bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Tumor-educated Gr1^{+}CD11b^{+} cells instigate breast cancer metastasis by twisting cancer cells plasticity via OSM/IL6–JAK signaling

Sanam Peyvandi[1, +, *], Manon Bulliard[1, +], Annamaria Kauzlaric[2], Oriana Coquoz[1], Yu-Ting Huang[1], Nathalie Duffey[1, +], Laetitia Gafner[1], Girieca Lorusso[1], Nadine Fournier[2], Qiang Lan[1,3,*,#], and Curzio Rüegg[1,*,#]

[1] Pathology Unit, Department of Oncology, Microbiology and Immunology (OMI), Faculty of Science and Medicine, University of Fribourg, 1700 Fribourg, Switzerland;

[2] Translational Data Science Group, Swiss Institute of Bioinformatics, Lausanne, Switzerland;

[3] Cell and Tissue Dynamics Research Program, Institute of Biotechnology, Helsinki Institute of Life Science (HiLIFE), University of Helsinki, Helsinki, Finland.

* These authors contributed equally as corresponding authors.

+ Current affiliations: S.P., Immunobiology Department, University of Lausanne (UNIL), Switzerland; M.B., Swiss Institute for Experimental Cancer Research (ISREC), Ecole Polytechnique Fédérale Lausanne (EPFL) UPCDA SV 2833 (Bâtiment SV) Lausanne, Switzerland; N.D., Department of Oncology, Centre Hospitalier Universitaire Vaudois (CHUV), University of Lausanne (UNIL), Switzerland; Y-T.H., Daiichi Sankyo Co., Ltd. Taipei City, Taiwan

Lead contacts and co-senior authors:

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

C.R. (curzio.ruegg@unifr.ch; +4126 300 8766), Department of Oncology, Microbiology, Immunology, Faculty of Science and Medicine, University of Fribourg, Chemin du Musée 18, Per 17, CH-1700 Fribourg, Switzerland

Q.L. (qiang.lan@helsinki.fi) Cell and Tissue Dynamics Research Program, Institute of Biotechnology, Helsinki Institute of Life Science (HiLIFE), Biocenter 2, P.O. Box 56, FIN-00014, University of Helsinki, Helsinki, Finland

Running title: OMS/IL6 twist breast cancer plasticity to promote metastasis

Keywords: Breast cancer metastasis, chemotherapy, cancer stem cells, GR1, Oncostatin M, IL6, JAK, plasticity, heterogeneity

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

Cancer cell plasticity contributes to tumor therapy resistance and metastasis formation, which represent the main causes of cancer-related death for most cancers, including breast cancer. The tumor microenvironment drives cancer cell plasticity and metastasis and unravelling the underlying cues may provide novel effective strategies to manage metastatic disease. Here we show that stem cell antigen-1 positive (Sca-1^{+}) murine breast cancer cells enriched during tumor progression and metastasis have higher *in vitro* cancer stem cell-like properties, enhanced *in vivo* metastatic ability, and initiate primary tumors rich in Gr1^{high}CD11b^{+}Ly6C^{low} cells. In turn, tumor-educated Gr1^{+}CD11b^{+} (Tu-Gr1^{+}CD11b^{+}) cells rapidly and transiently convert low metastatic 4T1-Sca-1^{-} cells into highly metastatic 4T1-Sca-1^{+} cells via secreted OSM and IL6. Moreover, chemotherapy-resistant and highly metastatic 4T1-derived cells maintain high Sca-1^{+} frequency through cell autonomous IL6 production. Inhibition of OSM, IL6 or JAK suppressed Tu-Gr1^{+}CD11b^{+}-induced Sca-1^{+} population enrichment *in vitro*, while JAK inhibition abrogated metastasis of chemotherapy-enriched Sca-1^{+} cells *in vivo*. Importantly, Tu-Gr1^{+}CD11b^{+} cells invoked a gene signature in tumor cells predicting shorter OS and RFS in breast cancer patients. Collectively, our data identified OSM/IL6-JAK as a clinically relevant paracrine/autocrine axis instigating breast cancer cell plasticity triggering metastasis.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Introduction

Metastasis accounts for over 90% of cancer-related death, calling for new strategies to prevent cancer cell dissemination and metastasis formation [1]. Recent studies using single-cell lineage tracing and single-cell RNA sequencing (scRNA-seq) technologies have provided detailed information about intratumor heterogeneity [2,3], whereby genetically, epigenetically and functionally diverse subpopulations of cancer cells exist within the tumor mass, spatially and temporally [4]. Intratumor heterogeneity may arise by modulating cancer cell plasticity, especially of cancer stem cells (CSCs), through cell-intrinsic and -extrinsic mechanisms [5,6].

Multiple CSC subpopulations appear to co-exist within the primary tumor mass resulting in high degree of tumor cell heterogeneity and increased aggressiveness [7] including in breast cancer [4,8–11]. In particular, CSCs can acquire metastasis initiating capacities [12] and resistance to therapy, resulting in cancer relapse [4]. Moreover, non-CSCs within the tumor bulk may acquire CSC properties to repopulate the tumor[4]. While CSC are defined rather functionally by their ability to initiate tumors and metastasis in low number *in vivo*, several cell surface markers associated with CSC features have been reported, including CD44, CD24, Sca-1, CD61, CD49f were used to identify breast CSCs [13,14].

The interaction of tumor cells with the tumor microenvironment (TME) contributes to tumor cell plasticity and tumor heterogeneity [15,16]. Cells of the TME also promote tumor escape and metastasis through multiple mechanisms, including promotion of angiogenesis, cell survival, invasion, epithelial-mesenchymal transition (EMT), and immunosuppression [17–22]. Recently, they have been reported to instigate expansion of CSC with metastatic ability (also known as metastasis-initiating cells) in different cancers [17,23,24]. Thus, the TME dynamics is a key driver of cancer cell plasticity and heterogeneity promoting tumor growth, progression and metastasis [4]. Accurate characterization of the

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

regulation of tumor plasticity and heterogeneity by the TME may reveal novel opportunities for developing effective anti-metastatic therapies [8,14].

TME-derived Oncostatin M (OSM) has been shown to mediate tumor progression and CSC stemness by activating its receptor OSMR [25]. OSM belongs to the IL6 family of cytokines (including IL6 itself, IL11 and LIF) [26,27], whose members bind to dimeric receptors sharing a common subunit (gp130 or IL6ST) and activate JAK-STAT, RAS-MAPK and PI3K-AKT pathways [28,29]. Increased OSM or IL6 expression correlates with reduced survival in breast cancer patients [30,31]. OSM was shown to drive breast cancer progression and metastasis through direct effects on cancer cells, such as suppression of estrogen receptor (ER) expression [31] and promotion of EMT [25,32], and indirect effects via TME cells, in particular the reprogramming of tumor associated macrophages and fibroblasts [33–36].

Here, by assessing the metastatic evolution of murine triple-negative breast cancer (TNBC) models *in silico* and *in vivo*, we observed that the Sca-1$^+$ tumor cell subpopulation is enriched during tumor progression. We show that tumor-educated Gr1$^+$CD11b$^+$ cells (Tu-Gr1$^+$CD11b$^+$), but not naïve Gr1$^+$CD11b$^+$ cells from spleen (Spl-Gr1$^+$CD11b$^+$) or bone marrow (BM-Gr1$^+$CD11b$^+$) in tumor-bearing mice, modulate tumor plasticity via OSM/IL6-JAK signaling by rapidly and transiently converting 4T1-Sca-1$^-$ cells into 4T1-Sca-1$^+$ cells with high metastatic capacity. Prolonged exposure of 4T1 cells to chemotherapy stably enriched for metastatic Sca-1$^+$ cells via an autocrine IL6-JAK signaling loop. A short *in vitro* treatment of these chemo-resistant cells with the JAK inhibitor Ruxolitinib, suppressed their metastatic capacity. Importantly, Tu-Gr1$^+$CD11b$^+$ invoked a gene expression signature in 4T1 cells that predicted shorter overall survival (OS) and relapse-free survival (RFS) in breast cancer patients, reinforcing the clinical relevance of these findings.

Our results reveal a novel mechanism modulating tumor plasticity and triggering the emergence of cancer cells with enhanced metastatic capacity, through paracrine (Tu-

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Gr1$^+$CD11b$^+$-mediated) and cell autonomous (chemotherapy-induced) OSM/IL6-JAK dependent signaling. The OSM/IL6-JAK axis may be considered as a candidate of actionable clinical targets to impinge on metastatic progression and therapy resistance.

Results

The Sca-1$^+$ population is enriched during tumor progression and has higher *in vivo* metastatic capacity

To investigate tumor cell heterogeneity during tumor progression, we first examined the expression of the previously reported breast CSC markers CD24, CD44, CD61, Sca-1 and CD49f [13] in the publicly available RNA sequencing (RNAseq) dataset from Ross *et al.* encompassing several murine breast cancer models [37]. This dataset includes data from cultured cancer cells (In_Culture), orthotopic primary tumors (OT_PT), spontaneous lung metastases (OT_LuM), and experimental lung metastases after tail vein injection (TV_LuM) (Fig. 1a and Supplementary Fig. 1a). *Sca-1* expression was elevated in lung metastasis in 4T1, 6DT1, Mvt1 and Met1 models compared with the respective primary tumors. Interestingly, in 4T1, 6DT1 and Mvt1 models, *Sca-1* expression was also elevated in experimental lung metastases compared with cultured cells. However, the expression of *Cd24*, *Cd44*, *Cd61* and *Cd49f* were not altered, or their expression pattern was inconsistent during progression (Supplementary Fig. 1a). Thus, increased *Sca-1* expression during metastasis is consistently observed in different preclinical breast cancer models.

To investigate whether the increased *Sca-1* expression within the tumor mass was due to an increased gene expression in all cancer cells or to the enrichment of a Sca-1$^+$ population, we orthotopically injected 4T1 tumor cells and determined the frequency of different cell populations present in the primary tumor and lung metastases 30 days later by flow cytometry (Fig. 1b). We observed that the frequency of both Sca-1$^+$ and CD61$^+$

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

populations increased in lung metastases compared to primary tumors (Fig. 1c). In contrast, the CD24$^+$, CD44$^+$ and CD49f$^+$ populations were not significantly altered.

The enrichment of the Sca-1$^+$ population in lung metastases prompted us to ask whether Sca-1$^+$ cells actively contribute to the metastasis. To this end, we isolated 4T1-Sca-1$^+$ and 4T1-Sca-1$^-$ cells by magnetic activated cell sorting (MACS) from parental 4T1 cells, which contains low frequency of Sca-1$^+$ population (10-15%) (Supplementary Fig. 1b, c), and examined their metastatic ability *in vivo*. In the orthotopic injection model, 4T1-Sca-1$^+$ cells formed significantly more lung metastases than 4T1-Sca-1$^-$ cells, while there was no significant difference in primary tumor growth (Fig. 1d-f). Upon tail vein injection, 4T1-Sca-1$^+$ cells displayed significantly greater lung colonization ability compared to 4T1-Sca-1$^-$ cells and a non-significant increase compared to parental 4T1 cells (Fig. 1g-i). In addition, 4T1-Sca-1$^+$ cells showed significantly higher *in vitro* mammosphere forming efficiency than 4T1-Sca-1$^-$ cells (Supplementary Fig. 2a), while *in vitro* cell growth, anchorage-independent survival, and cell motility were comparable (Supplementary Fig. 2b-d).

These results suggest that 4T1-Sca-1$^+$ and 4T1-Sca-1$^-$ cells have similar tumorigenic potential, while 4T1-Sca-1$^+$ cells have higher metastasis-initiating capacity.

The Sca-1$^+$ population is plastic *in vitro* and *in vivo*

Growing evidence indicates that cancer cells possess plastic features, which can be modulated by both cell-intrinsic factors and microenvironmental cues [37,38]. To characterize the observed plasticity of 4T1-Sca-1$^+$ cells, we first investigated isolated 4T1-Sca-1$^+$ and 4T1-Sca-1$^-$ cells *in vitro*. The abundance of Sca-1$^+$ population in 4T1-Sca-1$^+$ cells enriched by MACS sorting (> 75%) gradually decreased to 50% after 4 days of culture (Supplementary Fig. 2e, upper panel), while the Sca-1$^-$ cells (negatively enriched by

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

MACS) regenerated a Sca-1$^+$ population (from less that 1% to 19%) (Supplementary Fig. 2e, lower panel). Consistently, after orthotopic injection of 4T1-Sca-1$^+$ cells, the abundance of the Sca-1$^+$ population in the derived tumors decreased from 75% to about 40% after 23 days of growth, while in tumors generated from 4T1-Sca-1$^-$ cells it increased from less than 1% to 15% (Supplementary Fig. 2f), similar to the frequency of Sca-1$^+$ population in tumors derived from parental 4T1 cells (Fig. 1c). In addition, when tumor cells derived from primary tumors and lung metastases of parental 4T1-injected mice were cultured *ex vivo*, the abundance of the Sca-1$^+$ population significantly decreased from 20% to 4.5% and 60% to 19%, respectively (Supplementary Fig. 2g).

From these observations we conclude that both Sca-1$^+$ and Sca-1$^-$ populations are highly plastic and this plasticity appears to be modulated *in vivo*.

Tumor-educated Gr1$^+$CD11b$^+$ cells expand the metastatic Sca-1$^+$ population

Immune cells in the TME are critical determinants of tumor cells functions and behaviors, including metastatic capacity [39]. To collect evidence for a potential correlation between immune cells and the Sca-1$^+$ population, we characterized the inflammatory cells infiltrating the orthotopic primary tumors. We observed a significant increase of the Gr1^{high}CD11b$^+$Ly6C^{low} population and a significant decrease of the Gr1^{low}CD11b$^+$Ly6C^{high} population in tumors derived from the 4T1-Sca-1$^+$ cells, compared to tumors derived from the 4T1-Sca-1$^-$ cells (Fig. 2a). Gr1^{high}CD11b$^+$Ly6C^{low} cells are immature myeloid progenitors mobilized from the bone marrow by tumor-derived signals capable of establishing an immunosuppressive environment facilitating tumor progression and metastasis [40,41]. Gr1$^+$CD11b$^+$ cells form a homogenous population in the circulation (Supplementary Fig. 3a), consistent with the literature [42], while in the TME they differentiated into two distinct subpopulations, Gr1^{high} and Gr1^{low}. To examine the direct

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

contribution of the tumor-educated Gr1$^+$CD11b$^+$ cells in promoting the enrichment of the Sca-1$^+$ population, we isolated Gr1$^+$ cells from primary tumors, bone marrow and spleen of 4T1 tumor-bearing BALB/c mice by MACS, and co-cultured them *in vitro* with parental 4T1 cells (Fig. 2b). Gr1$^+$ cells isolated from tumors (Tu-Gr1$^+$CD11b$^+$), but not from spleen (Spl-Gr1$^+$CD11b$^+$) or bone marrow (BM-Gr1$^+$CD11b$^+$), significantly induced the expansion of a Sca-1$^+$ population in 4T1 cells (from 12.5% to 81%, $p<0.001$) in 48 hours (Fig. 2c). To test whether the expansion of a Sca-1$^+$ population from 4T1 cells required direct contact with Tu-Gr1$^+$CD11b$^+$ or was mediated by soluble factors, we compared the induction in two different co-culture setups, either in standard wells (cell contact) or in Transwells, where Tu-Gr1$^+$CD11b$^+$ and 4T1 cells were separated by a filter with 0.4 µm pores (Fig. 2d). We did not observe any significant difference in the induction efficiency of Sca-1$^+$ populations (measured by flow cytometry) between the two conditions and increasing the 4T1:Tu-Gr1$^+$CD11b$^+$ cell ratio from 1:1 to 1:3 did not further expand the Sca-1$^+$ population (Fig. 2e). Furthermore, conditioned medium from co-cultured Tu-Gr1$^+$CD11b$^+$ and 4T1 cells was also capable of subsequently expanding the Sca-1$^+$ population from 4T1 cells alone (Supplementary Fig. 3b). Interestingly, the Tu-Gr1$^+$CD11b$^+$ induced Sca-1$^+$ population appeared more stable in time compared to the isolated 4T1-Sca-1$^+$ cells (Supplementary Fig. 3c and Supplementary Fig. 2e). More importantly, when Gr1$^+$CD11b$^+$ cells-primed 4T1 cells were injected into the tail vein, Tu-Gr1$^+$CD11b$^+$ primed ones showed higher lung colonization capacity compared to Spl-Gr1$^+$CD11b$^+$ primed one (Fig. 2f-h).

These results imply that tumor-educated Gr1$^+$CD11b$^+$ cells induce the emergence of a highly metastatic Sca-1$^+$ population through secreted factors.

Tu-Gr1$^+$CD11b$^+$-induced and tumor-inherent Sca-1$^+$ populations display distinct gene expression profiles

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

To unravel the molecular basis for the metastatic capacity of the inherent 4T1- Sca-1$^+$ population and the Tu-Gr1$^+$CD11b$^+$ induced Sca-1$^+$ population, we first performed transcriptomic profiling of 4T1-Sca-1$^+$ and 4T1-Sca-1$^-$ cells isolated from the parental 4T1 line. Pathway enrichment analysis showed that 4T1-Sca-1$^+$ and 4T1-Sca-1$^-$ cells expressed different genes associated with distinct signaling pathways (Fig. 3a). The top 200 significantly upregulated and downregulated genes were extracted as Sca1 Positive and Sca1 Negative signatures, respectively (Supplementary Table 1). Next, we performed transcriptomic profiling of Tu-Gr1$^+$CD11b$^+$-primed 4T1, Spl-Gr1$^+$CD11b$^+$-primed 4T1 and parental 4T1 cells. Pathway enrichment analysis revealed that Tu-Gr1$^+$CD11b$^+$ and Spl-Gr1$^+$CD11b$^+$ priming induced distinct transcriptomic alterations in 4T1 cells (Fig. 3b). To focus on the transcriptomic alternations related to the Sca-1$^+$ population conversion, we compared Tu-Gr1$^+$CD11b$^+$-primed vs Spl-Gr1$^+$CD11b$^+$-primed cells. Interestingly, Tu-Gr1$^+$CD11b$^+$-primed 4T1 cells expressed both Sca-1 Positive and Sca-1 Negative signatures (Fig. 3c), suggesting that Tu-Gr1$^+$CD11b$^+$ may twist tumor plasticity by converting the 4T1-Sca-1$^-$ cells into 4T1-Sca-1$^+$ cells, rather than expanding the pre-existing 4T1-Sca-1$^+$ population.

To test whether, despite the diverse transcriptional profiles across different Sca-1$^+$ populations, there could be a common molecular mechanism underlying their induction and metastatic capacity, we compared the significantly differentially expressed genes between 4T1-Sca-1$^+$ versus 4T1-Sca-1$^-$ and Tu-Gr1$^+$CD11b$^+$ primed 4T1 versus Spl-Gr1$^+$CD11b$^+$ primed 4T1 cells. Strikingly, among a total of 1118 up- and 423 down-regulated genes found when comparing the two conditions, only 56 up- and one down-regulated genes were shared (Fig. 3d and Supplementary Table 1). This observation was consistent with the notion that the Sca-1$^+$ population in Tu-Gr1$^+$CD11b$^+$ primed 4T1 cells was different from the inherent 4T1-Sca-1$^+$ cells. Nonetheless, the fact that the 4T1-Sca-

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1$^+$ cells and Tu-Gr1$^+$CD11b$^+$ primed 4T1 cells possessed similar *in vivo* metastatic capacity suggested that among these common pathways some were relevant for 4T1 metastases formation. To this end, we analyzed the publicly available scRNA-seq dataset from 4T1 primary tumors of Sebastian et al. [43]. In this dataset, several cell types, including cancer cells, epithelial cells, fibroblasts, distinct subpopulations of myeloid cells were identified [43]. To determine significant ligand-receptor interactions from the scRNA-seq data, we performed cell-cell interaction analysis with CellPhoneDB [44] by focusing on the interactions between epithelial/cancer cells and myeloid cells (Fig. 3e, f). The analysis identified 160 ligand-receptor interaction pairs (Supplementary Table 2). Among those pairs, OSM receptor (OSMR) and Pyrimidinergic receptor P2Y6 (P2RY6) were the only ones present inside the 56 common genes shown in Fig. 3d. However, P2RY6 interacts with COPA (Coatomer Complex Subunit Alpha), a membrane protein involved in membrane traffic between endoplasmic reticulum and Golgi [45] and, thus, unlikely to mediate cell-cell contact-independent induction of Sca-1$^+$ population. As the IL6-JAK-STAT3 signaling pathway was upregulated both in the 4T1-Sca-1$^+$ population and Tu-Gr1$^+$CD11b$^+$-primed 4T1 cells (Fig. 3a, b), we next examined the expression of *Osm*, *Osmr*, *Il6st*, *Il6*, and *Il6* receptor (*Il6ra*) in the Sebastian dataset (Fig. 3g). The expression of *Il6* and *Osm* was restricted to myeloid cells, with *Osm* expression being more prominent, similar to a previous report [33]. *Osmr* was predominantly expressed in tumor cells, while *Il6st* and *Il6ra* were homogenously expressed in all cell types.

We then explored their expression in the 4T1-Sca-1$^+$ cells and Tu-Gr1$^+$CD11b$^+$ primed 4T1 cells. *Osm* and *Il6* expression were very low in all samples (normalized count number less than 7 on average) (Fig. 3h, i), consistent with data in the Sebastian dataset (Fig. 3g). *Osmr* and *Il6ra*, however, were highly expressed in 4T1-Sca-1$^+$ cells compared with 4T1-Sca-1$^-$ cells, while the expression of *Il6st* was abundant in both populations,

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

though higher in 4T1-Sca-1$^+$ cells (Fig. 3h). On the other hand, only *Osmr* was significantly upregulated in Tu-Gr1$^+$CD11b$^+$ primed 4T1 compared with Spl-Gr1$^+$CD11b$^+$ primed 4T1 (Fig. 3i).

Taken together, these results suggest that OSM/OMSR and IL6/IL6R signaling pathways may be involved in the Tu-Gr1$^+$CD11b$^+$-mediated expansion of 4T1-Sca-1$^+$ cells.

Tu-Gr1$^+$CD11b$^+$ cells promote Sca-1$^-$ to Sca-1$^+$ population conversion

The above results strongly implied that Tu-Gr1$^+$CD11b$^+$ cells convert Sca-1$^-$ population into the Sca-1$^+$ one. To further test this hypothesis, we compared the cell population dynamics in cultured cells and the orthotopic primary tumor by analyzing publicly available scRNA-seq datasets. By integrating scRNA-seq data from 3D cultured 4T1 cells (GSM4812003)[46] and tumor cells isolated from orthotopically fat pad-injected 4T1 primary tumor (PT) (GSM3502134)[47] (Fig. 4a, b), we observed 5 clusters. Clusters 0, 1, 2, 4 were predominant in cultured tumor cells, while cluster 3 was predominant in primary tumors. Single-cell trajectories analysis confirmed that cluster 3 was at the end of the transformation process (Fig. 4c, d). The population dynamics also showed that the fraction of cells in clusters 1, 2 and 4 decreased during the transformation, the one in cluster 0 only minimally increased, while the fraction in cluster 3 massively increased (Fig. 4e). Importantly, very few cultured 4T1 cells expressed *Sca-1* while it was abundantly expressed in the majority of cells in the primary tumor (Supplementary Fig. 4a, upper panel). Consistently, the fraction of cells expressing *Osmr* was higher in the primary tumor compared to cultured cells (Supplementary Fig. 4a, lower panel). Similar observations were obtained when analyzing scRNA-seq data from ER$^+$ human breast cancer model MCF-7. After integrating data from cultured MCF-7 cells (GSM4681765) and tumor cells that were isolated from MCF-7 intraductal injected mammary gland (GSM5904917)[48], 6

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

clusters were identified (Fig. 4f), with clusters 1 and 3 predominant in the cultured MCF-7 cells, while clusters 2 and 4 were predominant in the primary tumor (Fig. 4g). Further analysis showed that clusters 2 and 4 expanded during *in vitro* to *in vivo* tumor cell transformation and represented nearly 50% of the *in vivo* primary tumor cells (Fig. 4h-j). Although there is no human homolog of *Sca1* gene, *OSMR* expressing cells were increased upon tumor implantation, especially in clusters 2 and 1 (Supplementary Fig.4b).

To further investigate the signals involved in this transformation, we performed GSAE analysis for cluster 3 in 4T1 cells and clusters 2 and 4 in MCF-7 cells, respectively (Fig. 4k-n). By comparing the Hallmark gene signatures, IL6-JAK-STAT3 signature was significantly upregulated in both cell populations (Fig. 4k, m). Interestingly, Sca1 Positive and Sca1 Negative signatures were both upregulated (Fig. 4l, n), which is consistent with our *ex vivo* induction experiment (Fig. 3c). To validate the involvement of Tu-Gr1$^+$CD11b$^+$ during the cell population transformation, we extracted the top 50 upregulated genes (Supplementary Table 3) identified by comparing the Tu-Gr1$^+$CD11b$^+$ with the Sp-Gr1$^+$CD11b$^+$-stimulated 4T1 cells as Tu-Gr1$^+$CD11b$^+$ induced signature. Both cell populations predominant in the primary tumor in both 4T1 and MCF-7 models, upregulated the Tu-Gr1$^+$CD11b$^+$ induced signature (Fig. 4l, n).

These data, together with our *in vivo* observations (Fig. 1a-c) and *ex vivo* coculture experiments (Fig. 2b-e and Fig. 3c), indicate that Tu-Gr1$^+$CD11b$^+$ convert the Sca-1$^-$ population to Sca-1$^+$ population, likely, via OSM/IL6 signaling pathway.

OSM/IL6-JAK pathway mediates Tu-Gr1$^+$CD11b$^+$-induced Sca-1$^+$ population enrichment

To experimentally interrogate the role of OSM/IL6 in modulating the Sca-1$^+$ population, we first measured *Osm* and *Il6* mRNA expression in Spl-Gr1$^+$CD11b$^+$ and Tu-

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Gr1$^+$CD11b$^+$. Indeed, both *Osm* and *Il6* mRNA levels were significantly elevated in Tu-Gr1$^+$CD11b$^+$ (Fig. 5a). To functionally validate the role of OSM/IL6 in the generation of Sca-1$^+$ population, we treated 4T1 cells directly with recombinant OSM and IL6 proteins *in vitro* and measured the effect on the Sca-1$^+$ population. After 2 days of treatment, both OSM and IL6 significantly increased the frequency of Sca-1$^+$ cells from 16% to 38.8% and 26.5%, respectively (Fig. 5b). Conversely, blocking OSM and IL6 activities using anti- OSM or -IL6 neutralizing antibodies significantly reduced the emergence of Sca-1$^+$ population induced by Tu-Gr1$^+$CD11b$^+$ conditioned medium (Fig. 5c). The combination of anti-OSM and -IL6 antibodies did not have additive effects suggesting that OSM and IL6 both contribute in promoting Sca-1$^+$ population by sharing the same signaling cascades.

OSMR and IL6R signal by activating the intracellular Janus tyrosine kinase (JAK) [49]. To explore the involvement of the JAK pathway in the emergence of the Sca-1$^+$ population, we treated 4T1 cells with the JAK inhibitor (Ruxolitinib) during exposure to recombinant OSM and IL6. Ruxolitinib treatment prevented the emergence of the Sca-1$^+$ population in response to recombinant OSM and IL6 (Fig. 5d).

From these results, we conclude that OSM/IL6-JAK pathway mediates Tu-Gr1$^+$CD11b$^+$ -induced Sca-1$^+$ population enrichment.

Tu-Gr1$^+$CD11b$^+$-induced Sca-1$^+$ population and 4T1-inherent Sca-1$^+$ population have distinct CSC and EMT gene expression profiles

OSM/IL6-JAK signaling has been reported to support tumor progression by promoting a CSC phenotype and epithelial-mesenchymal plasticity [25,36,50]. To further characterize CSC and EMT features in 4T1-Sca-1$^+$ cells and Tu-Gr1$^+$CD11b$^+$ induced Sca-1$^+$ population, we used our RNAseq data to explore the expression of 17 stem cell and EMT markers (Supplementary Fig. 5). The stem cell markers *Oct4* (*Pou5f1*), *Sox2* and *Nanog*

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

were undetectable or very low in all or some samples. 4T1-Sca-1$^+$ cells had higher expression of *Aldh1a1*, *Aldh3a1* and *Podxl* but lower expression of *Klf4* and *Sox9* compared to 4T1-Sca-1$^-$ cells. There was no difference in the expression of *Abcg2* and *Has2*. Tu-Gr1$^+$CD11b$^+$ primed 4T1 cells had higher expression of *Klf4* and *Has2*, lower expression of *Aldh1a1*, *Aldh3a1* and *Sox9*, and similar expression of *Podxl* when compared with Spl-Gr1$^+$CD11b$^+$ primed 4T1. Among them, only *Has2* expression was specifically elevated in Tu-Gr1$^+$CD11b$^+$ primed 4T1 compared with control 4T1 and Spl-Gr1$^+$CD11b$^+$-primed 4T1 (Supplementary Fig. 5a). On the other hand, 4T1-Sca-1$^+$ cells had lower expression of *Cdh1* and higher expression of *Snail1*, *Twist1*, *Vim* and *Foxc1* which support an EMT status, although *Zeb1* expression was reduced (Supplementary Fig. 5b). Globally, the expression of most of the EMT genes were similar between Tu-Gr1$^+$CD11b$^+$ and Spl-Gr1$^+$CD11b$^+$ primed 4T1 cells, except for *Snail2* and *Vim*, whose expression was suppressed in Tu-Gr1$^+$CD11b$^+$ primed 4T1 (Supplementary Fig. 5b).

Taken together, these results indicate that the Tu-Gr1$^+$CD11b$^+$-induced Sca-1$^+$ population and inherent Sca-1$^+$ population have different CSC and EMT transcriptional profiles, reinforcing the notion that, Tu-Gr1$^+$CD11b$^+$-induced Sca-1$^+$ population are not just enriched tumor-inherent Sca-1$^+$ population

Chemotherapy enriches a Sca-1$^+$ population with CSC features

CSC and cancer cell plasticity contribute to drug resistance in various tumor types, including breast cancer [51–55]. The above results, including the significantly elevated expression of *Aldh3a1* (a marker of drug resistance) in the 4T1-Sca-1$^+$ population prompted us to investigate the resistance of this population to chemotherapy. To this end, we treated 4T1 cells for 48 hours *in vitro* with methotrexate (MTX) and doxorubicin (Dox), two widely used chemotherapy drugs, including in breast cancer. The 48 hours treatment

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

of either drug increased the frequency of Sca-1$^+$ population in 4T1 cells (Fig. 6a). Next, we mimicked a clinically relevant situation of cancer cells escaping chemotherapy by exposing 4T1 cells to 28 nM MTX, which is slightly higher than the IC_{50} concentration of the drug (REF!!), for up to 3 weeks and recovered the surviving cells by switching to normal medium (Fig. 6b). The selected cell line, named MR13, was highly enriched in Sca-1$^+$ cells (>60%) (Fig. 6c). Compared to parental 4T1 tumor cells, MR13 cells exhibited a higher mammosphere forming efficiency (Fig. 6d), lower *in vitro* proliferative capacity (Fig. 6e), increased survival under non-adhesive conditions (Fig. 6f) and increased cell mobility (Fig. 6g), which were consistent with CSC-like properties. When tested in a 48-hours cytotoxicity assay, MR13 cells were more resistant against MTX compared to parental 4T1 (Supplementary Fig. 6a).

MR13-derived tumors are highly metastatic and rich in Gr1^{high}CD11b$^+$ Ly6C^{low} cells

To characterize the *in vivo* behavior of MR13 cells, we orthotopically implanted them into BALB/c mice and monitored the progression. MR13 cells formed smaller primary tumors relative to parental 4T1 cells, that were more metastatic to the lung (Fig. 6h-j) and enriched in Gr1^{high}CD11b$^+$ Ly6C^{low} cells compared to 4T1 tumors (Fig. 6k), similarly to 4T1-Sca-1$^+$-derived tumors (Fig. 2a). Strikingly, we observed metastases in the heart (Supplementary Fig. 6b), which we never observed with the parental 4T1 cells. MR13 cells retained a large fraction of the Sca-1$^+$ population *in vitro*, even when cultured in the absence of MTX (Fig. 6c) and upon *in vivo* expansion (Fig. 6l). Such stability of Sca-1 population contrasted with 4T1-Sca-1$^+$ cells and Tu-Gr1$^+$CD11b$^+$-primed 4T1 cells (Supplementary Fig. 2e, g) that, upon isolation or induction reverted to a Sca-1$^-$ phenotype, suggesting that MR13 line was capable of self-sustain its own Sca-1$^+$ population.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Taken together, chemotherapy-selected MR13 cells share some similar *in vivo* characteristics as 4T1-Sca-1$^+$ cells, while they are capable of self-sustaining high Sca-1$^+$ abundancy both *in vitro* and *in vivo.*

IL6/IL6R-JAK autocrine signaling maintains Sca-1 positivity and metastatic capacity in MR13 cells

To better understand the chemotherapy-induced alterations in MR13 cells, we performed transcriptomic analyses comparing MR13 and parental 4T1 cells. Pathway enrichment analysis showed that the IL6-JAK-STAT3 signature was also elevated in MR13 cells (Fig. 7a). Importantly, MR13 gene expression significantly positively correlated with the Sca1 Positive signature. At the same time, it negatively correlated with the Sca1 Negative signature (Fig. 7b). This observation suggested that chemotherapy enriched for the inherent 4T1-Sca-1$^+$ population, rather than converting plastic Sca-1$^-$ cells, as observed upon Tu-Gr1$^+$CD11b$^+$-stimulation. In addition, *Osmr*, *Il6*, *Il6ra* and *Il6st* were all overexpressed in MR13 cells compared to parental 4T1 cells, while the *Osm* expression was not altered (Fig. 7c). To functionally validate the IL6-JAK signaling pathway in Sca-1$^+$ population maintenance, we treated MR13 cells with Ruxolitinib *in vitro* for 48 hours. The treatment significantly decreased the fraction of Sca-1$^+$ cells (Fig. 7d). Importantly, *in vitro* treatment of MR13 cells with Ruxolitinib for 3 days nearly completely abolished their lung metastatic capacity upon tail vein injection (Fig. 7e-f).

Altogether, these data suggest that MR13 cells sustain the metastatic Sca-1$^+$ population through cell-autonomous activation of the IL6-JAK signaling pathway.

Tu-Gr1$^+$CD11b$^+$ invoked tumor cell signature predicts shorter overall and relapse-free survival in breast cancer patients

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

To evaluate the clinical relevance of the crosstalk between Tu-Gr1$^+$CD11b$^+$ and tumor cells, we tested whether the Tu-Gr1$^+$CD11b$^+$-induced 4T1 signature could predict cancer progression in patients. To this end, we interrogated the METABRIC dataset [56] with the Tu-Gr1$^+$CD11b$^+$-induced signature. Thirty-two human orthologue genes (Supplementary Table 3) in the murine 50 genes signature were present in the METABRIC dataset. Patients with the higher expression level of the signature had shorter overall survival (OS; p = 0.0056) and relapse-free survival (RFS; p = 0.032) (Fig. 8a, b). Notably, OSM expression positively correlated with the signature in all patients (Supplementary Fig. 7a), suggesting OSM do contribute to altered signature expression in patients. Of the thirty-two genes, five (*MX1, IRF7, OAS1, CMPK2, ISG15*) were commonly discriminant for a shorter OS and RFS (Fig. 8c, d), and taken together, they further enhanced the predictive power (OS: p=0.00055; RFS: p=0.00069) (Fig. 8e, f).

These data indicate that a small set of genes issued from the murine tumor cell signature invoked by Tu-Gr1$^+$CD11b$^+$ cells can predict a shorter OS and RFS in breast cancer patients, thereby reinforcing the clinical significance of the proposed model.

Discussion (1470 words -> 1689 NOW … OMG)

Metastatic disease and therapy resistance are the leading causes of breast cancer mortality, calling for novel approaches to effectively prevent and cure for metastasis and therapy resistance. This is particularly relevant for TNBC, where in spite of the recent encouraging results with targeted therapies, such as PARP inhibitors for tumors with germline BRCA mutations, or checkpoint inhibitors for PD-L1$^+$ tumors, management of metastatic disease remains challenging [57–60]. It has been proposed that CSCs present in the primary tumor are responsible for tumor persistence, metastasis, and therapy resistance [52,61]. Enrichment and differentiation of CSC contribute to tumor heterogeneity

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[17,19–22]. Importantly, CSC features can be intrinsic or plastic and be modulated by cues from the TME [62–64].

Here, we have interrogated the contribution of the Sca-1$^+$ population to breast cancer metastasis and its modulation by the TME. We reported that tumor-educated Gr1$^+$CD11b$^+$ cells (Tu-Gr1$^+$CD11b$^+$) instigate cancer metastasis by twisting cancer cell plasticity and enriching for a Sca-1$^+$ population, with enhanced metastatic capacity. We identified OSM/IL6-JAK as a paracrine communication axis between Tu-Gr1$^+$CD11b$^+$ and breast cancer cells and as an autocrine loop in chemotherapy-resistant tumor cells, promoting tumor heterogeneity, CSC features and metastatic capacity. Importantly, breast cancer patients expressing high levels of the human orthologues of the gene expression signatures invoked by Tu-Gr1$^+$CD11b$^+$ have significantly shorter OS and RFS, reinforcing the clinical significance of our findings. While some of these elements have been reported individually before, our results extend these observations by providing an integrative view of paracrine (Tu-Gr1$^+$CD11b$^+$-induced) and autocrine (chemotherapy-induced) communication in regulating tumor heterogeneity, cancer cell plasticity, metastasis, and resistance to chemotherapy.

A main observation stemming from this study is that Sca-1$^+$ population exists under three different conditions: as inherent 4T1-Sca-1$^+$ population, upon exposure to Tu-Gr1$^+$CD11b$^+$ cells and upon chemotherapy treatment (MR13). All these populations had higher metastatic ability compared to their counterpart controls. GSEA analysis suggested that the Tu-Gr1$^+$CD11b$^+$-induced Sca-1$^+$ population was likely to convert from the Sca-1$^-$ population (Fig. 3c and Fig. 4), while the Sca-1$^+$ population surviving chemotherapy (MR13) appeared to be enriched from an inherent Sca-1$^+$ population (Fig. 7b). In addition, the different gene expression signature of Tu-Gr1$^+$CD11b$^+$-induced Sca-1$^+$ population relative to the inherent Sca-1$^+$ cells suggest a remarkable functional

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

plasticity of these cells. This plasticity is further supported by comparing single cell gene expression of murine 4T1 and human MCF-7 breast cancer cells, using publicly available scRNA-seq datasets, before and after *in vivo* growth (Fig. 4). *In vivo*, tumor cells undergo a transformation enriching for Sca-1 (in the mouse), OSMR expressing cells, Sca1 Positive, Sca1 Negative and Tu-Gr1^{+}CD11b^{+}-induced signatures, and IL6-JAK signaling pathway from precursor cells (Sca-1^{-} in the mouse) (Fig. 4). Gong *et al.* have reported that sorted Sca-1^{-} 4T1 cells could be transiently transformed into a Sca-1^{+} population by radiotherapy [65]. An analogous observation was reported in colorectal cancer, where selective ablation of LGR5^{+} CSCs in organoids leads to initial tumor regression, followed by regrowth driven by LGR5^{+} CSCs reemerging from the LGR5^{-} population [66]. Taken together, our observations further consolidate the notion that cancer consists of a heterogenous and plastic tumor mass, including highly metastatic cell populations, by demonstrating that tumor recruited and educated Gr1^{+}CD11b^{+} cells contribute to such plasticity by inducing the conversion of low metastatic Sca-1^{-} population into a highly metastatic Sca-1^{+} population. Dedicated time course scRNA-seq analyses together with lineage tracing experiments may help to further characterize the detailed origin, development, fate, and function of these Sca-1^{+} populations during cancer progression.

Recruitment and accumulation of Gr1^{+}CD11b^{+} cells in the TME, particularly through the chemokines CCL2, CXCL1 and CXCL2, or IL-33, is considered a critical step for their contribution to tumor progression and metastasis [67–70]. Consistent with these observations, tumors derived from sorted 4T1-Sca-1^{+} cells, or MR13 cells that are intrinsically enriched for Sca-1^{+} cells, have a higher content of Gr1^{high}CD11b^{+}Ly6C^{low} cells compared to tumors derived from 4T1-Sca-1^{-} or parental 4T1 cells, respectively (Fig. 2a and 5k). Beyond recruitment, tumor-mediated education of Gr1^{+}CD11b^{+} cells appears to be necessary to gain higher metastatic activity: Only Gr1^{+}CD11b^{+} cells recovered from primary tumors (but

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

not from spleen or bone marrow) induced Sca-1 positivity and enhanced metastatic ability in 4T1 cells (Fig. 2c, e). The role of Gr1$^+$CD11b$^+$ cells in promoting metastasis has been mainly attributed to promotion of angiogenesis, EMT and immunosuppression [17,68]. Peng *et al.*, reported the ability of those cells to endow CSC-like features to breast cancer cells but their metastatic capacity was not interrogated [17]. Our observations suggest a self-sustaining positive feedback mechanism between highly metastatic cancer cells (Sca-1$^+$ population) and Gr1$^+$CD11b$^+$ cells: inherent Sca1$^+$ CSC-like, metastatic, cells promote recruitment and local education of Gr1$^+$CD11b$^+$ cells, which in turn promote tumor heterogeneity, cancer cell plasticity and metastatic capacity by converting low-metastatic Sca1$^-$ cells into additional high-metastatic Sca1$^+$ cells (Fig. 9).. Recently, it was reported that neutrophils escorting blood circulating tumor cells (CTCs) expands the metastatic potential of CTCs [71]. While this effect was attributed to the promotion of cell cycle progression of CTCs through direct contact with the neutrophils, in light of our findings, one may also consider the possibility that clustered neutrophils may also promote the expansion of a CSCs-like phenotype with higher metastatic capacity. Interestingly, OSM was reported to be expressed by neutrophils cocultured with breast cancer cells [36] and to promote phenotypic changes associated with mesenchymal and stem cell-like differentiation in breast cancer [36,72]. Together with our observation, these findings further reinforce the notion that boosting the *Osm* expression in Gr1$^+$CD11b$^+$ cells is part of their educating program prompted by the tumor. One outstanding question raised by these observations is by which mechanisms and pathways cancer cells, in particular Sca-1$^+$ ones, educate Gr1$^+$CD11b$^+$ cells to acquire cancer plasticity-promoting activity.

Besides directly activating tumor cells, OSM has also been shown to remodel macrophages and fibroblasts of the TME [33,36]. Araujo *et al.* recently reported that OSM derived from tumor-infiltrating myeloid cells reprogram fibroblasts to secrete VEGF and the

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

chemokines CXCL1 and CXCL16, resulting in enhanced myeloid cell recruitment and breast cancer progression [33]. Here we extend these observations by demonstrating that tumor-educated Gr1$^+$CD11b$^+$-derived OSM/IL6 twist cancer cell plasticity by promoting a rapid but reversible conversion of Sca-1$^-$ cells into more metastatic Sca-1$^+$ cells (Fig. 5b). We further broaden the implications of the OSM/IL6-JAK axis, by demonstrating that MR13 cells that have escaped chemotherapy hijacked this paracrine mechanism in a cell autonomous manner by elevating the *Il6* expression (Fig. 9). Significantly, a short *in vitro* treatment with Ruxolitinib effectively abrogated their metastatic capacity (Fig. 7e, f).

While OSM/OSMR is the only interaction pair identified by the cell-cell interaction analysis (Fig. 3f, g), it is possible that other molecular mechanisms may also play roles in modulating Sca-1$^+$ cell plasticity and metastasis. One candidate is the IL6/IL6R communication axis, which was less prominent in the cell-cell communication analysis (Fig. 3f, g), but was highly expressed in our experimental data in Tu-Gr1$^+$CD11b$^+$ compared with Spl-Gr1$^+$CD11b$^+$ (Fig. 5a), and in chemotherapy-resistant MR13 cells compared with 4T1 (Fig. 7c). Moreover, the neutralization of IL6 also suppressed the Tu-Gr1$^+$CD11b$^+$ conditioned medium induced Sca-1$^+$ cell enrichment, much alike OSM inhibition (Fig. 5c).

Importantly, we demonstrated that a human orthologue signature of the 4T1 gene expression signature invoked by Tu-Gr1$^+$CD11b$^+$ can predict a significantly shorter OS and RFS in breast cancer patients (Fig. 8). This finding strengthens the clinical significance of the observed crosstalk between Tu-Gr1$^+$CD11b$^+$ and tumor cells. Strikingly, the five genes that significantly contribute to the discriminatory power of the signature are genes related to native or viral immunity or regulated by interferon. While expression of interferons and interferon response genes in breast cancer has been mainly associated with tumor suppression and improved survival [73] , there is evidence also correlating interferon responses with tumor promotion, therapy resistance and reduced survival [74]. As

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

JAKs/STATS are activated by both IFN and OSM/IL6 receptors, it is conceivable that OSM/IL6 only activates a subset of the IFN-induced genes with tumor-promoting activity, as the case for Mx1 [75]. Consistent with our findings, JAK/STAT signaling has recently been shown to initiate the lineage plasticity in prostate cancer as well as to promote lineage plasticity-driven targeted therapy resistance in a stem-like subpopulation of prostate cancer [76,77]. On the other hand, Aouad *et al.* showed that epithelial-mesenchymal plasticity is essential for the generation of a dormant cell state of ER^+ breast cancer during progression, and the activation of IL6-JAK-STAT signaling triggers tumor cell awakening and recurrence [48].

One crucial question raised by these results is whether the OSM/IL6-JAK pathway is a potential actionable clinical target to impinge on metastatic progression and therapy resistance. In particular, the observation that a short *in vitro* treatment of the highly metastatic MR13 cells profoundly suppressed their metastatic capacity *in vivo* (Fig. 7e, f), suggests potential long-lasting effects consistent with an adjuvant effect. JAK inhibitors are being tested in clinical trials in breast cancer. A phase I study combining Ruxolitinib with paclitaxel in HER2-negative metastatic breast cancer showed good tolerability and evidence of activity [78]. A phase I/II trial of Ruxolitinib in combination with trastuzumab in metastatic HER2 positive breast cancer and a phase II study combining Ruxolitinib with capecitabine in advanced $HER2^-$ breast cancer, however, did not improve progression-free survival [79,80]. The absence of benefits in these studies in advanced breast cancer and our reported mechanistic observations on metastatic progression, raise the question of whether the JAK inhibitors should be considered in adjuvant setting in high-risk patients, to prevent progression to metastases, rather than treating patients already bearing metastases.

In conclusion, we reported here that a subpopulation of tumor cells within the tumor

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

mass educates requited $Gr1^+CD11b^+$ cells to convert a low metastatic subpopulation into highly metastatic one, through the OSM/IL6-JAK signaling axis (Fig. 9). The clinical relevance of this observation is supported by human transcriptomic data. This process is hijacked by tumor cells that survived chemotherapy and evolved toward a highly metastatic phenotype via cell autonomous IL6-JAK signaling. Importantly, a short *in vitro* treatment with a clinically approved JAK inhibitor, Ruxolitinib, suppresses their metastatic capacity these cells *in vivo.* These results should stimulate considering testing JAK inhibitors in the adjuvant setting in TNBC breast cancer patients at high-risk for metastatic progression.

Methods

Cell culture. The 4T1 murine breast cancer cell line was kindly provided by Dr Fred R. Miller (Michigan Cancer Foundation, Detroit, MI, USA). 4T1 cells were cultured in high glucose DMEM supplemented with 10% heat-inactivated FBS, 1% penicillin–streptomycin (P/S, from Gibco) and 1% Non-Essential Amino Acid (Gibco).

Tumor models. 4T1, MR13, sorted 4T1-Sca-1^+ and 4T1-Sca-1^- ($5x10^4$ cells in 50µl PBS/10% of 8.1 mg/ml Matrigel Matrix, were injected in the fourth right mammary gland of mice. Prior to surgery, ketamine (1.5 mg/kg) and xylazine (150 mg/kg) (both from Graeub) were injected intra-peritoneally to anesthetize the animals. Immune cell populations were analyzed at different time points post-tumor cell injection. Tumor length and width were measured twice a week with caliper and used to calculate tumor volume by the following equation: volume = (length x $width^2$) x π/6. Tumors were collected and weighted at necropsy. For the intravenous injections, $2x10^5$ sorted 4T1-Sca-1^+ and 4T1- Sca-1^- tumor cells resuspended in a volume of 50 µl of PBS were injected into the mice tail vein. Lung

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

metastases were quantified 10 days post-injection. At each indicated time point, mice were sacrificed according to defined ethical criteria and were killed by CO_2 inhalation followed by neck dislocation or terminal bleeding. All animal procedures were performed in accordance with the Swiss legislations on animal experimentation and approved by the Cantonal Veterinary Service of the Cantons Vaud and Fribourg for experiments in Lausanne and Fribourg (VD_1486.2; 2017_34_FR; 2017_34_FR, 2014_58_FR; 2011-33-FR).

Reagents and chemicals. Growth factor reduced Matrigel Matrix (MG) was obtained from Becton Dickinson (BD Biosciences). Collagenase I was purchased from Worthington and DNAse I from Roche. Bovine serum albumin (BSA), crystal violet (CV) and paraformaldehyde (PFA) were obtained from Sigma-Aldrich. Drugs, inhibitors and cytokines: Doxorubicin and Methotrexate (generously provided by the Department of Oncology, University Hospital, University of Lausanne, Lausanne, Switzerland), Ruxolitinib (JAK inhibitor, Cat N°7064, Biotechne), anti-mouse Oncostatin M (R&D systems), anti-mouse IL-6 (BioxCell), recombinant mouse Oncostatin M and IL-6 (Biolegend, Cat #: 762802 & 575702 respectively).

Antibodies. The following anti-mouse antibodies were used following manufacturer's instructions: anti-CD16/CD32 Fc blocking antibody (BD Biosciences), anti-CD24-FITC (clone M1/69, eBioscience), anti-CD29-PE and-PE-Cy5 (clone HMβ1-1, BioLegend), anti-Sca-1-APC (clone D7, eBioscience), anti-CD61-Alexa 647 (clone 2C9.62(HMβ3-1), BioLegend, anti-CD45-PE (clone 30-F11, BD Biosciences), anti-Gr1-eFlour450 (clone RB6-8C5,eBioscience), anti-Ly6C FITC (clone HK1.4, BioLegend), anti-Ly6G-APC (clone 1A8-Ly6g,eBioscience), anti-F4/80 PerCP/Cy5 (clone BM8, eBioscience), anti-CD11b-PE-

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Cy7 (clone M1/70,eBioscience), anti-CD11c-APC-eFluor780 (clone N418, eBioscience), anti-CD4-FITC (clone 6K1.5, eBioscience), anti-CD8-PE (clone 53-6.7, eBioscience), anti-B220-APC (clone RA3-6B2, eBioscience), anti-CD49b-eFlour450 (clone DX5, BioLegend), Annexin V-APC (clone B217656, BioLegend), Propidium Iodide-PerCP (Clone V13245, Life technologies).

Magnetic cell sorting (MACS). MACS separators were used for positive and negative cell selections based on manufacturer's instructions. Briefly, cells were counted and resuspend in 500 µl of MACS buffer with 10 µl of fluorescent coupled antibody of interest (APC-conjugated anti-Sca-1 and PE-conjugated anti-Gr1) per 10^7 cells added. Cells were incubated for 30 minutes in the dark at 4°C and then washed with MACS buffer. After centrifugation and resuspension in 80 µl of MACS buffer per 10^7 cells, 20 µl of anti- APC-conjugated magnetic beads (Miltenyi Biotec) per 10^7 cells was added. After a 20 minutes incubation in the dark at 4°C, cells were washed with MACS buffer. After the centrifugation they were resuspended in MACS buffer and the magnetic separation was performed using LS MACS column (maximum 10^8 labeled cells) for positive selection and LD MACS column (maximum 10^8 labeled cells) for negative selection (Miltenyi Biotec). The purity of positive subpopulation was >70% and <99% for the negative subpopulation.

Co-culture of $Gr1^+CD11b^+$ cells sorted from tumors or spleen of 4T1 tumor bearing mice. $Gr1^+$ cells were sorted from tumors or spleen at day 23 post-injection (see below Magnetic beads cell sorting section). Afterwards, $1.5x10^5$ 4T1 cells were co-cultured in 6 wells plates using Transwell plates (0.4 µm, Nunc, ThermoFisher Scientific) with $Gr1^+CD11b^+$ sorted cells (top) and 4T1 cells (bottom) at different 4T1: $Gr1^+CD11b^+$ ratios

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(1:1, 1:3 and 1:5). After 48 hours of co-culture, cells were analyzed for Sca-1 expression and gene expression by flow cytometry and semi-quantitative real-time qPCR as indicated.

Flow cytometry analysis on tissue samples. Mice were sacrificed at different time points for blood and tumors collection. Tumors were cut in small pieces with scissors, washed, and digested in serum free medium supplemented with Collagenase I and DNAse I (Roche). The mixture was incubated at 37 °C for 45 minutes on a shaking platform. Subsequently, serum-supplemented medium was added to neutralize the enzymatic reaction and the tissue suspensions were filtered through a 100 µm and a 70 µm sterile nylon gauzes. Upon centrifugation (5 minutes at 1400 rpm), pellets were recovered and red blood cells lysed with ACK buffer (Biolegend). The staining procedure and the flow cytometry acquisition are as described previously [81]. Data acquisition was performed using the FACSCalibur (BD Biosciences) or MACSQuant flow cytometer from Miltenyi Biotec and data analyzed by FlowJo v10.0.7 (tree Stat Inc.).

***In vitro* cell proliferation assay**. Cells were collected and seeded in tissue culture 96-wells-plates (Costar) at 3’000 cells/well. Cells were grown in complete medium for 24, 36, 72 and 96 hours. At each time point cells were washed once with PBS, then fixed with 4% PFA and stained with 0.5% crystal violet solution for 0.5 hours. The stained cells were gently washed with deionized water to remove the extra dye and air-dried overnight at room temperature. After solubilizing the dye with crystal violet eluting buffer (70% ethanol and 1% acetic acid), cell viability was assessed by reading the absorbance at 595 nm wavelength in a multiwell plate reader (Modulus II microplate reader, Turner Biosystems). Results were analyzed by Prism (Graph pad software, Inc.) expressed as mean values of optical density (OD) of octuplet determinations ± SEM.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

***In vitro* cytotoxic assay.** Tumor cells were plated at a concentration of 3'000 cell/well into 96-wells plates. The following day, a series of concentrations of the different drugs were supplemented to the culture medium. Untreated control cells were kept in normal culture medium. Cell viability of each well was assessed with crystal violet staining 48 hours after treatment, as described above. Results were analyzed by Prism software by a non-linear regression analysis and expressed as relative cell viability compared with non-treated control. The 50% maximum inhibition concentrations (IC_{50}) were used to determine the drug-resistant ability of treated cells.

Real-time reverse transcription qPCR and primers. Changes in mRNA expression levels were determined by semi-quantitative real-time qPCR. RNA samples were obtained from adherent cells using RNeasy kit from QIAGEN according to manufacturer's instructions. From each sample, 1 µg RNA was retro-transcribed using SuperScript II Reverse Transcriptase kit (Life Technologies – Invitrogen), according to manufacturer's instructions. The reactions were performed in a StepOnePlus™ thermocycler (Applied Biosystems, Life Technologies – Invitrogen) using the KapaSYBR® FAST SYBR Green Master Mix (Kapa Biosystems). Each reaction was performed in triplicates and values were normalized to murine 36B4 housekeeping gene. The comparative C_t method was used to calculate the difference of gene expression between samples. The following murine-specific primers (Microsynth AG) were used:
Sca-1 (F:5'-TCAGGAGGCAGCAGTTATTGTG-3',R: 5'-TGGCAACAGGAAGTCTTCACG-3'), 36B4(F:5'-GTGTGTCTGCAGATCGGGTAC-3',R:5'-CAGATGGATCAGCCAGGAAG-3'), OSM (F: 5'- ATGCAGACACGGCTTCTAAGA-3', R: 5'- TTGGAGCAGCCACGATTG G-3'), OSMR: (F: 5'- CATCCCGAAGCGAAGTCTTGG-3', R: 5'-

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

GGCTGGGACAGTCCATTC TAAA-3'), IL-6: (F: 5'- TACCACTTCACAAGTCGGAGGC-3', R: 5'- CTGCAAGTGCATCAT CGTTGTTC-3').

Histopathology. Tumors and lungs were harvested at the end of the experiments, fixed in formalin and embedded in paraffin. 5 µm thick serial sections were cut from the tissue blocks. 3-4 sections taken at 100 µm distance were stained with hematoxylin and eosin (H&E) and used to assess tumor morphology and quantify lung metastasis. Slides were scanned by Nanozoomer (Hamamatsu Photonics) and metastasis were counted manually using NDP.viewer2 software (Hamamatsu Photonics). Metastatic index was calculated by normalizing the metastasis number with the volume of primary tumor.

Anoikis assay. A 6 wells plate was coated with 500 µl of 3% Poly (2-hydroxyethyl methacrylate) (poly-HEMA) (Sigma-Aldrich) dissolved in 95% ethanol overnight. Wells without poly-HEMA served as control. The day after, $2x10^5$ cells resuspended in DMEM medium were seeded. The plate was incubated at 37°C for 24 hours. Cells were collected, centrifuged and resuspended in 100 µl of 1x -binding buffer (1 ml of 5x annexin-binding buffer (Life Technologies, Invitrogen) in 4 ml of deionized water). 3 µl APC annexin V (BioLegend) and 1 µl of 100 µg/ml Propidium Iodide (PI) working solution (5 µl of 1mg/ml PI stock, Life Technologies, Invitrogen) in 45 µl of 1x annexin-binding buffer) were added to each 100 µl cell suspension. Cells were incubated for 15 minutes at room temperature protected from light. 200 µl of 1x annexin V-binding buffer was added to the cells and samples were kept on ice. Apoptotic (Annexin V^{pos} and PI^{neg}) and dead (PI^{pos}) cells were analyzed by flow cytometry. Results were expressed as percentage of dead cells ± SEM.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Mammosphere-forming assay 5'000 cells/well were seeded in non-adhesive U bottom 96-wells plate in a semi-solid MEGM medium supplemented with 20 ng/ml EGF, and 20 ng/ml bFGF and heparin. Medium was gently replaced every 3 to 4 days. After 11 to 14 days culture, mammospheres of 50 to 150 µm diameter were detected under the microscope (bright field) and counted to quantify the sphere formation efficiency (SFE) as percentage of the initial number of seeded cells per well.

Bulk RNA sequencing and data analysis. Four independent sorts of 4T1-Sca-1^{+} and 4T1-Sca-1^{-} cells (as described in the magnetic cell sorting section) or four independent 4T1 cells primed with Tu- Gr1^{+}CD11b^{+} or Spl-Gr1^{+}CD11b^{+} were prepared. RNAs of these cells were isolated using the NucleoSpin RNA protocol of Macherey-Nagel (as described in real-time (RT) qPCR and primers section). Samples were normalized for 1 µg RNA in a volume of 20 µl, sequenced on the NextSeq500 sequencer using the NextSeq 500/550 HT reagent v2 kit (Illumina) at the Swiss Integrative Center for Human Health (SICHH) in Fribourg, or the Lausanne Genomics Technologies Facility (GTF, UNIL) in Lausanne, Switzerland. For data analysis, all sequencing reads were processed for quality control, removal of low quality reads, adaptor sequence and ribosomal RNA by fastqc(0.11.8) [82], multiqc (1.9) [83], Trimmomatic (0.39) [84] and SortMeRNA(2.1) [85] accordingly. The filtered reads were mapped to the reference genome (mm10) using htseq-count (0.6.1) [86] or Salmon (0.99.0) [87]. The normalization of the read counts and the analysis of the differential expression between the groups of samples were performed with in R(v4.1.3), a free software environment available at https://www.r-project.org/ using packages DESeq2 (v3.15) [89]. Pathway enrichment analysis was performed using packages GSVA(1.42.0) [90] and GSEABase(1.56.0) [91] with the input of DESeq2 normalized count numbers using ssgsea method comparing the Hallmark genesets from MSigDB (v7.4.1) [92] with default settings.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The significant altered pathways were determined by computing moderated t-statistics and false discovery rates with the limma(3.50.3)[93] for pair-wised comparison. The heatmaps were produced with R package pheatmap(1.0.12)[94] with default settings while pathway hierarchy clustering was performed by similarity based on Euclidean distance and the ward aggregation algorithm. The Sca1 Positive and Sca1 Negative signatures were extracted from the top 200 most upregulated or downregulated genes in 4T1-Sca-1^{+} and 4T1-Sca-1^{-} RNAseq data, respectively, with the threshold of adjusted p-value <0.05, fold change >1.5 or <-1.5 and average normalized count number >20. For Venn diagram, the genes fulfilling the threshold of adjusted p-value < 0.05, fold change > 1.5 or < -1.5 and average normalized count number >20 are compared. The figures were produced with R package venn (1.10) [95] . Further analysis and figures generation were performed in R using packages tidyverse(1.3.1)[96], ggplot2(3.3.6) [97], circlize(0.4.15) [98], biomaRt(2.50.3) [99,100], RColorBrewer (1.1-3) [101], clusterProfiler (4.2.2) [102,103], enrichplot (1.14.2) [104], ggpubr(v0.4.0)[105] ggbreak (0.1.0) [105].

Microarray hybridization and data analysis Experiments were performed as previously described [106]. Briefly, triplicates wells of cultured 4T1 and MR13 cells were used for RNA extraction using RNeasy kit (QIAGEN). Probe synthesis and GeneChip Mouse Gene Exon 1.0 ST Array (Affymetrix Ltd) hybridization were performed at the GTF, UNIL, Lausanne, Switzerland. Microarray analyses were carried out with R. After quantification of gene expression with robust multi-array normalization [100] using the BioConductor package Affy, (http://www.bioconductor.org/) significance of differential gene expression was determined by computing moderated t-statistics and false discovery rates with the limma package [93]. Annotation was based on the genome version NCBI Build 36 (Feb. 2006). The obtained p-values were corrected for multiple testing by calculating estimated false discovery rates

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(FDR) using the method of Benjamini-Hochberg. Heatmaps were produced by color-coding gene-wise standardized log gene expression levels (mean zero standard deviation one). Probe-sets were shown hierarchically clustered by similarity based on Euclidean distance and the ward aggregation algorithm.

Public RNAseq data analysis The RPKM normalized gene expression data in Ross dataset (GSE150928) from multiple murine models of breast cancer metastasis was obtained from Gene Expression Omnibus (GEO) in the NCBI data repository. The data were analyzed and plot with R package tidyverse(1.3.1)[96], ggplot2 (3.3.6) [97] and ggpubr (v0.4.0) [107]. The single-cell RNAseq data in Sebastian dataset [43] was obtained from Dryad data repository (https://doi.org/10.6071/M3238R). The data were analyzed with R package Seurat (3.0) [108]. The tumor and different myeloid cell populations were extracted for cell-cell interaction analysis using CellPhondDB (2.0) [44]. The identified interaction pairs were extracted and plot using circlize (0.4.15) [98] and ComplexHeatmap (2.11.2) [109]. The ligands and receptors annotated in the circular plot were compiled from databases in CellTalkDB[110], SingleCellSignalR [111]. To investigate the tumor cell dynamics, datasets for 4T1 (GSE158844 and GSM3502134) and MCF-7 (GSM4681765 and GSM5904917) were obtained from GEO. The data were filtered and normalized separately before merging with IntegrateData function included in Seurat. Cell cycle regression was performed according to the standard protocol of Seurat. For MCF-7, due the huge difference of the sample size, 2000 cells were randomly selected from GSM4681765 data, and then merged with GSM5904917. Single-cell trajectories analysis was performed with monocle 3 [112–115] with default settings and the root and start point were selected manually for pseudotime caculation. The GSEA analysis of selected clusters was performed with R package fgsea[116] and the gene rank was calculated with Wilcoxon rank sum test and auROC

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

analysis using wilcoxauc function included in presto package. The Tu-Gr1$^+$CD11b$^+$-induced signature was extracted by comparing gene expression between Tu-Gr1$^+$CD11b$^+$ and Spl-Gr1$^+$CD11b$^+$-educated 4T1 cells with adjusted p-value <0.05, fold change >2, and the top 50 genes were selected.

Clinical data analysis. To validate our finding in clinical data, the human orthologs of murine Tu-Gr1$^+$CD11b$^+$-induced signature genes were used. Conversion from murine to human gene symbols and Entrez IDs was performed with the biomaRt package (2.46.3) [99,100], using the reference mart https://dec2021.archive.ensembl.org. Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) breast cancer data was downloaded from cBioPortal [117–119] in August 2022, and expression data was log2 transformed. Expression values were stratified in two groups by median values. Survival curves were generated using the ggsurvplot function from the survminer package (0.4.9)[107], and were compared between groups using a log-rank test. Survival curves were created using the survfit function from the survival package (3.2.11) [120]. Cox proportional hazard regression model was performed through the coxph function of the same package.

Statistical analyses. Unless specified, the data were presented as mean ± SEM from at least 3 independent experiments, unless otherwise indicated. Statistical comparisons were performed by an unpaired Student's t test with a two-tailed distribution or one-way ANOVA analysis of variance with Bonferroni post-test, using Prism 7.0 GraphPad Software, Inc.

Graphic illustrations. Illustrative schemes were created with BioRender.com.

Data and code availability

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The raw and processed bulk RNAseq data used to generate figures in Fig. 3 and Supplementary Fig. 5 and microarray data used to generate figures in Fig. 7 have been deposited in the GEO database under the access code GSEXXXXXX.

The code used for the analyses is open-source and available through with R packages described in methods.

References

1. Fares, J., Fares, M. Y., Khachfe, H. H., Salhab, H. A. & Fares, Y. Molecular principles of metastasis: a hallmark of cancer revisited. *Signal Transduct. Target. Ther.* **5**, 1–17 (2020).
2. Valdés-Mora, F. *et al.* Single-cell transcriptomics reveals involution mimicry during the specification of the basal breast cancer subtype. *Cell Rep.* **35**, 108945 (2021).
3. Yang, D. *et al.* Lineage tracing reveals the phylodynamics, plasticity, and paths of tumor evolution. *Cell* **185**, 1905-1923.e25 (2022).
4. Torborg, S. R., Li, Z., Chan, J. E. & Tammela, T. Cellular and molecular mechanisms of plasticity in cancer. *Trends Cancer* **8**, 735–746 (2022).
5. Baccelli, I. & Trumpp, A. The evolving concept of cancer and metastasis stem cells. *J. Cell Biol.* **198**, 281–293 (2012).
6. Lawson, D. A., Kessenbrock, K., Davis, R. T., Pervolarakis, N. & Werb, Z. Tumour heterogeneity and metastasis at single-cell resolution. *Nat. Cell Biol.* **20**, 1349–1360 (2018).
7. Celià-Terrassa, T. & Jolly, M. K. Cancer Stem Cells and Epithelial-to-Mesenchymal Transition in Cancer Metastasis. *Cold Spring Harb. Perspect. Med.* **10**, a036905 (2020).
8. Dagogo-Jack, I. & Shaw, A. T. Tumour heterogeneity and resistance to cancer therapies. *Nat. Rev. Clin. Oncol.* **15**, 81–94 (2018).
9. McDonald, K.-A. *et al.* Tumor Heterogeneity Correlates with Less Immune Response and Worse Survival in Breast Cancer Patients. *Ann. Surg. Oncol.* **26**, 2191–2199 (2019).
10. Hata, A. N. *et al.* Tumor cells can follow distinct evolutionary paths to become resistant to epidermal growth factor receptor inhibition. *Nat. Med.* **22**, 262–269 (2016).

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

11. Ramirez, M. *et al.* Diverse drug-resistance mechanisms can emerge from drug-tolerant cancer persister cells. *Nat. Commun.* **7**, 10690 (2016).

12. Oskarsson, T., Batlle, E. & Massagué, J. Metastatic stem cells: sources, niches, and vital pathways. *Cell Stem Cell* **14**, 306–321 (2014).

13. Ye, F. *et al.* CD49f Can Act as a Biomarker for Local or Distant Recurrence in Breast Cancer. *J. Breast Cancer* **20**, 142–149 (2017).

14. Vaillant, F. *et al.* The Mammary Progenitor Marker CD61/β3 Integrin Identifies Cancer Stem Cells in Mouse Models of Mammary Tumorigenesis. *Cancer Res.* **68**, 7711–7717 (2008).

15. Li, J. J., Tsang, J. Y. & Tse, G. M. Tumor Microenvironment in Breast Cancer—Updates on Therapeutic Implications and Pathologic Assessment. *Cancers* **13**, 4233 (2021).

16. Ge, R., Wang, Z. & Cheng, L. Tumor microenvironment heterogeneity an important mediator of prostate cancer progression and therapeutic resistance. *Npj Precis. Oncol.* **6**, 1–8 (2022).

17. Peng, D. *et al.* Myeloid-Derived Suppressor Cells Endow Stem-like Qualities to Breast Cancer Cells through IL6/STAT3 and NO/NOTCH Cross-talk Signaling. *Cancer Res.* **76**, 3156–3165 (2016).

18. Li, K. *et al.* Myeloid-derived suppressor cells as immunosuppressive regulators and therapeutic targets in cancer. *Signal Transduct. Target. Ther.* **6**, 1–25 (2021).

19. Lorusso, G. & Ruegg, C. The tumor microenvironment and its contribution to tumor evolution toward metastasis. *Histochem Cell Biol* **130**, 1091–103 (2008).

20. Sleeman, J. P. *et al.* Concepts of metastasis in flux: the stromal progression model. *Semin Cancer Biol* **22**, 174–86 (2012).

21. Cha, Y. J. & Koo, J. S. Role of Tumor-Associated Myeloid Cells in Breast Cancer. *Cells* **9**, (2020).

22. Pastaki Khoshbin, A., Eskian, M., Keshavarz-Fathi, M. & Rezaei, N. Roles of Myeloid-Derived Suppressor Cells in Cancer Metastasis: Immunosuppression and Beyond. *Arch Immunol Ther Exp Warsz* **67**, 89–102 (2019).

23. Panni, R. Z. *et al.* Tumor-induced STAT3 activation in monocytic myeloid-derived suppressor cells enhances stemness and mesenchymal properties in human pancreatic cancer. *Cancer Immunol. Immunother.* **63**, 513–528 (2014).

24. Cui, T. X. *et al.* Myeloid-derived suppressor cells enhance stemness of cancer cells by inducing microRNA101 and suppressing the corepressor CtBP2. *Immunity* **39**,

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

611–621 (2013).

25. Abaurrea, A., Araujo, A. M. & Caffarel, M. M. The Role of the IL-6 Cytokine Family in Epithelial–Mesenchymal Plasticity in Cancer Progression. *Int. J. Mol. Sci.* **22**, 8334 (2021).

26. Zarling, J. M. *et al.* Oncostatin M: a growth regulator produced by differentiated histiocytic lymphoma cells. *Proc. Natl. Acad. Sci. U. S. A.* **83**, 9739–9743 (1986).

27. Felcher, C. M., Bogni, E. S. & Kordon, E. C. IL-6 Cytokine Family: A Putative Target for Breast Cancer Prevention and Treatment. *Int J Mol Sci* **23**, (2022).

28. Johnson, D. E., O'Keefe, R. A. & Grandis, J. R. Targeting the IL-6/JAK/STAT3 signalling axis in cancer. *Nat Rev Clin Oncol* **15**, 234–248 (2018).

29. Quintas-Cardama, A. & Verstovsek, S. Molecular pathways: Jak/STAT pathway: mutations, inhibitors, and resistance. *Clin Cancer Res* **19**, 1933–40 (2013).

30. Tawara, K. *et al.* HIGH expression of OSM and IL-6 are associated with decreased breast cancer survival: synergistic induction of IL-6 secretion by OSM and IL-1beta. *Oncotarget* **10**, 2068–2085 (2019).

31. West, N. R., Murphy, L. C. & Watson, P. H. Oncostatin M suppresses oestrogen receptor-α expression and is associated with poor outcome in human breast cancer. *Endocr. Relat. Cancer* **19**, 181–195 (2012).

32. Guo, L. *et al.* Stat3-coordinated Lin-28-let-7-HMGA2 and miR-200-ZEB1 circuits initiate and maintain oncostatin M-driven epithelial-mesenchymal transition. *Oncogene* **32**, 5272–5282 (2013).

33. Araujo, A. M. *et al.* Stromal oncostatin M cytokine promotes breast cancer progression by reprogramming the tumor microenvironment. *J Clin Invest* **132**, (2022).

34. Bryson, B. L., Junk, D. J., Cipriano, R. & Jackson, M. W. STAT3-mediated SMAD3 activation underlies Oncostatin M-induced Senescence. *Cell Cycle* **16**, 319–334 (2017).

35. Taniguchi, K. & Karin, M. IL-6 and related cytokines as the critical lynchpins between inflammation and cancer. *Semin. Immunol.* **26**, 54–74 (2014).

36. West, N. R., Murray, J. I. & Watson, P. H. Oncostatin-M promotes phenotypic changes associated with mesenchymal and stem cell-like differentiation in breast cancer. *Oncogene* **33**, 1485–1494 (2014).

37. Ross, C. *et al.* Metastasis-Specific Gene Expression in Autochthonous and Allograft Mouse Mammary Tumor Models: Stratification and Identification of Targetable Signatures. *Mol. Cancer Res. MCR* **18**, 1278–1289 (2020).

38. Fico, F., Bousquenaud, M., Rüegg, C. & Santamaria-Martínez, A. Breast Cancer

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Stem Cells with Tumor- versus Metastasis-Initiating Capacities Are Modulated by TGFBR1 Inhibition. *Stem Cell Rep.* **13**, 1–9 (2019).

39. Hanahan, D. Hallmarks of Cancer: New Dimensions. *Cancer Discov.* **12**, 31–46 (2022).

40. Ouzounova, M. *et al.* Monocytic and granulocytic myeloid derived suppressor cells differentially regulate spatiotemporal tumour plasticity during metastatic cascade. *Nat. Commun.* **8**, 14979 (2017).

41. Veglia, F., Sanseviero, E. & Gabrilovich, D. I. Myeloid-derived suppressor cells in the era of increasing myeloid cell diversity. *Nat. Rev. Immunol.* **21**, 485–498 (2021).

42. Bronte, V. *et al.* Recommendations for myeloid-derived suppressor cell nomenclature and characterization standards. *Nat. Commun.* **7**, 12150 (2016).

43. Sebastian, A. *et al.* Single-Cell Transcriptomic Analysis of Tumor-Derived Fibroblasts and Normal Tissue-Resident Fibroblasts Reveals Fibroblast Heterogeneity in Breast Cancer. *Cancers* **12**, 1307 (2020).

44. Efremova, M., Vento-Tormo, M., Teichmann, S. A. & Vento-Tormo, R. CellPhoneDB: inferring cell–cell communication from combined expression of multi-subunit ligand–receptor complexes. *Nat. Protoc.* **15**, 1484–1506 (2020).

45. Brandizzi, F. & Barlowe, C. Organization of the ER–Golgi interface for membrane traffic control. *Nat. Rev. Mol. Cell Biol.* **14**, 382–392 (2013).

46. Chen, K. *et al.* Phenotypically supervised single-cell sequencing parses within-cell-type heterogeneity. *iScience* **24**, 101991 (2021).

47. Yeo, S. K. *et al.* Single-cell RNA-sequencing reveals distinct patterns of cell state heterogeneity in mouse models of breast cancer. *eLife* **9**, e58810 (2020).

48. Aouad, P. *et al.* Epithelial-mesenchymal plasticity determines estrogen receptor positive breast cancer dormancy and epithelial reconversion drives recurrence. *Nat. Commun.* **13**, 4975 (2022).

49. Jin, W. Role of JAK/STAT3 Signaling in the Regulation of Metastasis, the Transition of Cancer Stem Cells, and Chemoresistance of Cancer by Epithelial–Mesenchymal Transition. *Cells* **9**, 217 (2020).

50. Junk, D. J. *et al.* Oncostatin M promotes cancer cell plasticity through cooperative STAT3-SMAD3 signaling. *Oncogene* **36**, 4001–4013 (2017).

51. Naz, F., Shi, M., Sajid, S., Yang, Z. & Yu, C. Cancer stem cells: a major culprit of intra-tumor heterogeneity. *Am. J. Cancer Res.* **11**, 5782–5811 (2021).

52. Zhang, R., Tu, J. & Liu, S. Novel molecular regulators of breast cancer stem cell

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

plasticity and heterogeneity. *Semin Cancer Biol* **82**, 11–25 (2022).

53. Qin, S. *et al.* Emerging role of tumor cell plasticity in modifying therapeutic response. *Signal Transduct. Target. Ther.* **5**, 228 (2020).

54. Kong, D., Hughes, C. J. & Ford, H. L. Cellular Plasticity in Breast Cancer Progression and Therapy. *Front. Mol. Biosci.* **7**, 72 (2020).

55. Muzio, G., Maggiora, M., Paiuzzi, E., Oraldi, M. & Canuto, R. A. Aldehyde dehydrogenases and cell proliferation. *Free Radic. Biol. Med.* **52**, 735–746 (2012).

56. Curtis, C. *et al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486**, 346–352 (2012).

57. Shen, M. *et al.* A review of current progress in triple-negative breast cancer therapy. *Open Med Wars* **15**, 1143–1149 (2020).

58. Al-Mahmood, S., Sapiezynski, J., Garbuzenko, O. B. & Minko, T. Metastatic and triple-negative breast cancer: challenges and treatment options. *Drug Deliv Transl Res* **8**, 1483–1507 (2018).

59. Azim, H. A., Ghosn, M., Oualla, K. & Kassem, L. Personalized treatment in metastatic triple-negative breast cancer: The outlook in 2020. *Breast J* **26**, 69–80 (2020).

60. Vagia, E., Mahalingam, D. & Cristofanilli, M. The Landscape of Targeted Therapies in TNBC. *Cancers Basel* **12**, (2020).

61. Hua, Z., White, J. & Zhou, J. Cancer stem cells in TNBC. *Semin Cancer Biol* **82**, 26–34 (2022).

62. Batlle, E. & Clevers, H. Cancer stem cells revisited. *Nat Med* **23**, 1124–1134 (2017).

63. Brooks, M. D., Burness, M. L. & Wicha, M. S. Therapeutic Implications of Cellular Heterogeneity and Plasticity in Breast Cancer. *Cell Stem Cell* **17**, 260–71 (2015).

64. Yeo, S. K., Wen, J., Chen, S. & Guan, J. L. Autophagy Differentially Regulates Distinct Breast Cancer Stem-like Cells in Murine Models via EGFR/Stat3 and Tgfbeta/Smad Signaling. *Cancer Res* **76**, 3397–410 (2016).

65. Gong, J. *et al.* Genotoxic stress induces Sca-1-expressing metastatic mammary cancer cells. *Mol. Oncol.* **12**, 1249–1263 (2018).

66. Shimokawa, M. *et al.* Visualization and targeting of LGR5+ human colon cancer stem cells. *Nature* **545**, 187–192 (2017).

67. Oo, M. W. *et al.* Resident stroma-secreted chemokine CCL2 governs myeloid-derived suppressor cells in the tumor microenvironment. *JCI Insight* **7**, (2022).

68. Franklin, D. A. *et al.* MEK activation modulates glycolysis and supports

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

suppressive myeloid cells in TNBC. *JCI Insight* **5**, (2020).

69. Bullock, K. & Richmond, A. Suppressing MDSC Recruitment to the Tumor Microenvironment by Antagonizing CXCR2 to Enhance the Efficacy of Immunotherapy. *Cancers Basel* **13**, (2021).

70. Jovanovic, I. P. *et al.* Interleukin-33/ST2 axis promotes breast cancer growth and metastases by facilitating intratumoral accumulation of immunosuppressive and innate lymphoid cells. *Int J Cancer* **134**, 1669–82 (2014).

71. Szczerba, B. M. *et al.* Neutrophils escort circulating tumour cells to enable cell cycle progression. *Nature* **566**, 553–557 (2019).

72. Doherty, M. R. *et al.* The opposing effects of interferon-beta and oncostatin-M as regulators of cancer stem cell plasticity in triple-negative breast cancer. *Breast Cancer Res* **21**, 54 (2019).

73. Cavalli, L. R., Riggins, R. B., Wang, A., Clarke, R. & Haddad, B. R. Frequent loss of heterozygosity at the interferon regulatory factor-1 gene locus in breast cancer. *Breast Cancer Res. Treat.* **121**, 227–231 (2010).

74. Weichselbaum, R. R. *et al.* An interferon-related gene signature for DNA damage resistance is a predictive marker for chemotherapy and radiation for breast cancer. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 18490–18495 (2008).

75. Aljohani, A. I. *et al.* Myxovirus resistance 1 (MX1) is an independent predictor of poor outcome in invasive breast cancer. *Breast Cancer Res. Treat.* **181**, 541–551 (2020).

76. Chan, J. M. *et al.* Lineage plasticity in prostate cancer depends on JAK/STAT inflammatory signaling. *Science* **377**, 1180–1191 (2022).

77. Deng, S. *et al.* Ectopic JAK–STAT activation enables the transition to a stem-like and multilineage state conferring AR-targeted therapy resistance. *Nat. Cancer* **3**, 1071–1087 (2022).

78. Lynce, F. *et al.* Phase I study of JAK1/2 inhibitor ruxolitinib with weekly paclitaxel for the treatment of HER2-negative metastatic breast cancer. *Cancer Chemother Pharmacol* **87**, 673–679 (2021).

79. Kearney, M. *et al.* Phase I/II trial of ruxolitinib in combination with trastuzumab in metastatic HER2 positive breast cancer. *Breast Cancer Res Treat* **189**, 177–185 (2021).

80. O'Shaughnessy, J. *et al.* A randomized, double-blind, phase 2 study of ruxolitinib or placebo in combination with capecitabine in patients with advanced HER2-negative breast cancer and elevated C-reactive protein, a marker of systemic inflammation. *Breast Cancer Res Treat* **170**, 547–557 (2018).

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

81. Lan, Q. *et al.* Type I interferon/IRF7 axis instigates chemotherapy-induced immunological dormancy in breast cancer. *Oncogene* **38**, 2814–2829 (2019).

82. FASTQC. A quality control tool for high throughput sequence data | BibSonomy. https://www.bibsonomy.org/bibtex/f230a919c34360709aa298734d63dca3.

83. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).

84. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

85. Kopylova, E., Noé, L. & Touzet, H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinforma. Oxf. Engl.* **28**, 3211–3217 (2012).

86. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).

87. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).

88. https://www.r-project.org/.

89. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

90. Hänzelmann, S., Castelo, R. & Guinney, J. GSVA: gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinformatics* **14**, 7 (2013).

91. Morgan M, Falcon S, Gentleman R. GSEABase: Gene set enrichment data structures and methods. R package version 1.58.0. (2022).

92. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci.* **102**, 15545–15550 (2005).

93. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).

94. Kolde, R. Pheatmap: pretty heatmaps. *R Package Version* **1**, 726 (2012).

95. Dusa, Adrian. Draw Venn Diagrams R package venn. (2016).

96. Wickham, H. *et al.* Welcome to the Tidyverse. *J. Open Source Softw.* **4**, 1686 (2019).

97. Wickham, H. *ggplot2*. (Springer, 2009). doi:10.1007/978-0-387-98141-3.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

98. Gu, Z., Gu, L., Eils, R., Schlesner, M. & Brors, B. circlize Implements and enhances circular visualization in R. *Bioinforma. Oxf. Engl.* **30**, 2811–2812 (2014).

99. Durinck, S. *et al.* BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinforma. Oxf. Engl.* **21**, 3439–3440 (2005).

100. Durinck, S., Spellman, P. T., Birney, E. & Huber, W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.* **4**, 1184–1191 (2009).

101. Neuwirth, E. RColorBrewer: ColorBrewer Palettes. 5.

102. Wu, T. *et al.* clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innov. Camb. Mass* **2**, 100141 (2021).

103. Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics J. Integr. Biol.* **16**, 284–287 (2012).

104. Yu, G. & Hu, E. enrichplot: Visualization of Functional Enrichment Result. (2022) doi:10.18129/B9.bioc.enrichplot.

105. Xu, S. *et al.* Use ggbreak to Effectively Utilize Plotting Space to Deal With Large Datasets and Outliers. *Front. Genet.* **12**, 774846 (2021).

106. Monnier, Y. *et al.* CYR61 and alphaVbeta5 integrin cooperate to promote invasion and metastasis of tumors growing in preirradiated stroma. *Cancer Res.* **68**, 7323–7331 (2008).

107. Kassambara, A. ggpubr: 'ggplot2' Based Publication Ready Plots. (2020).

108. Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e21 (2019).

109. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinforma. Oxf. Engl.* **32**, 2847–2849 (2016).

110. Shao, X. *et al.* CellTalkDB: a manually curated database of ligand–receptor interactions in humans and mice. *Brief. Bioinform.* **22**, bbaa269 (2021).

111. Cabello-Aguilar, S. *et al.* SingleCellSignalR: inference of intercellular networks from single-cell transcriptomics. *Nucleic Acids Res.* **48**, e55 (2020).

112. Trapnell, C. *et al.* The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014).

113. Qiu, X. *et al.* Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods* **14**, 979–982 (2017).

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

114. Cao, J. *et al.* The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496–502 (2019).

115. McInnes, L., Healy, J. & Melville, J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. Preprint at http://arxiv.org/abs/1802.03426 (2020).

116. Korotkevich, G. *et al.* Fast gene set enrichment analysis. 060012 Preprint at https://doi.org/10.1101/060012 (2021).

117. Gao, W. *et al.* IL20RA signaling enhances stemness and promotes the formation of an immunosuppressive microenvironment in breast cancer. *Theranostics* **11**, 2564–2580 (2021).

118. Cerami, E. *et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* **2**, 401–404 (2012).

119. Ramos, M. *et al.* Multiomic Integration of Public Oncology Databases in Bioconductor. *JCO Clin. Cancer Inform.* 958–971 (2020) doi:10.1200/CCI.19.00119.

120. Therneau, T. M. & Grambsch, P. M. *Modeling Survival Data: Extending the Cox Model*. (Springer, 2000). doi:10.1007/978-1-4757-3294-8.

Acknowledgments

The authors wish to thank Sarah Cattin and Melissa Rizzo for assistance with FACS analysis, Dr. Ana-Marija Sulić (Institute of Biotechnology, HiLIFE, University of Helsinki) for the insightful discussion about scRNA-seq data analysis, Dr. Fred R. Miller (Michigan Cancer Foundation, Detroit,MI, USA) for providing 4T1 cells, Dr. Khalil Zaman (Department of Oncology, University Hospital, University of Lausanne, Lausanne, Switzerland) for providing chemotherapeutic drugs. This work was supported by grants from the Swiss National Science Foundation (31003A_179248; 310030_208136), the Swiss Cancer League (KFS 4400-02-2018) and the Medic Foundation (to C.R.).

Author contributions

Conceptualization, S.P., Q.L., C.R.; Methodology, S.P., N.F., Q.L., C.R.; Software, A.K., N.F., Q.L.; Validation, S.P., M.B., A.K., Q.L.; Formal analysis, S.P., M.B., A.K., Q.L.;

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Investigation, S.P., M.B., A.K., O.C., Y-T.H., N.D., L.G., Q.L.; Resources, S.P., M.B., A.K., O.C., N.D., L.G., G.L.; Data Curation, S.P., M.B., A.K., N.D., Q.L., C.R.; Writing – original draft preparation, S.P., Q.L., C.R.; Writing – review and editing, S.P., M.B., A.K., Y-T.H., G.L., N.D., Q.L., C.R.; Visualization, S.P., M.B., A.K., Q.L., C.R.; Supervision, S.P., Q.L., C.R.; Project administration, Q.L., C.R.; Funding acquisition, C.R.

Competing interests

The authors have no conflicts of interest to declare.

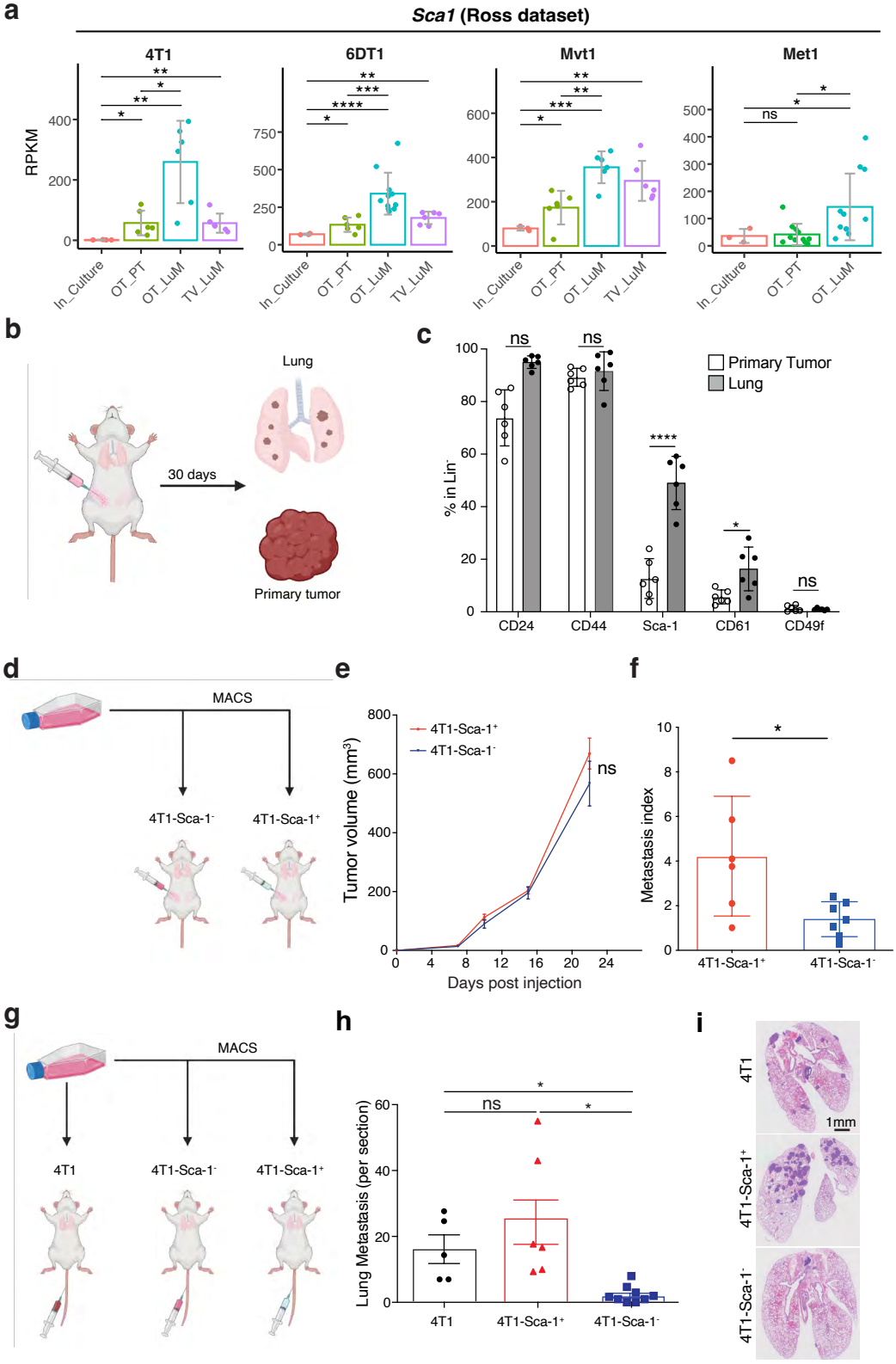

a
Sca1 (Ross dataset)
4T1
6DT1
Mvt1
Met1
RPKM
In_Culture
OT_PT
OT_LuM
TV_LuM
ns
b
Lung
30 days
Primary tumor
c
Primary Tumor
Lung
% in Lin-
CD24
CD44
Sca-1
CD61
CD49f
d
MACS
4T1-Sca-1-
4T1-Sca-1+
e
Tumor volume (mm3)
Days post injection
f
Metastasis index
g
4T1
h
Lung Metastasis (per section)
i
1mm

Figure 1

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 1. Sca-1$^+$ population is enriched during *in vivo* metastasis across multiple breast cancer models

(a) Sca-1 mRNA expression in the metastatic murine breast cancer models 4T1, 6DT1, Mvt1 and Met1, extracted from the Ross dataset. Analyzed samples consist of cultured cells (In_Culture), orthotopic injected primary tumors (OP_PT), spontaneous lung metastases (OP_LuM) and lung metastases induced by tail-vain injection (TV_LuM). Data are presented as mean of Reads Per Kilobase of transcript per Million mapped reads (RPKM) ± SD (unpaired two-tailed student's t test with Bonferroni correction).

(b) Experimental set up for *in vivo* experimental validation. The 4T1 tumor cells were orthotopically injected into the 4th mammary fat pad. Thirty days later, cells from primary tumors and lungs were isolated to examine CSC marker expression by flow cytometry.

(c) Frequency of CSC marker expression in primary tumors and lung metastases. Results give the percentage of CD24, CD44, Sca-1, CD61 and CD49f positive cells gated in lineage negative cells (CD45$^-$CD31$^-$TER119$^-$).

(d-f) Experimental set up **(d)** of the *in vivo* experiment to assess tumor growth **(e)** and lung metastatic ability (metastatic index) **(f)** of 4T1-Sca-1$^+$ and 4T1-Sca-1$^-$ populations isolated from parental 4T1 cells orthotopically injected into the 4th mammary fat pad. Metastases are assessed 21 days after tumor cell injection. n=8/group, 3 independent experiments.

(g-i) Experimental set up **(g)** of the *in vivo* experiment to assess lung colonization capacity of sorted parental 4T1, 4T1-Sca-1$^+$ and 4T1-Sca-1$^-$ cells by tail-vain injection. Lung metastatic nodule numbers **(h)** and representative images **(i)** of lungs from mice 10 days post injection (n=5-6, 2 independent experiments). Scale bar=1 mm. Data are

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

represented as mean values ± SEM (unpaired two-tailed student's t test). P values: *, $p<0.05$; **, $p< 0.01$; ***, $p< 0.001$; ****, $p< 0.0001$; ns, non-signifiant.

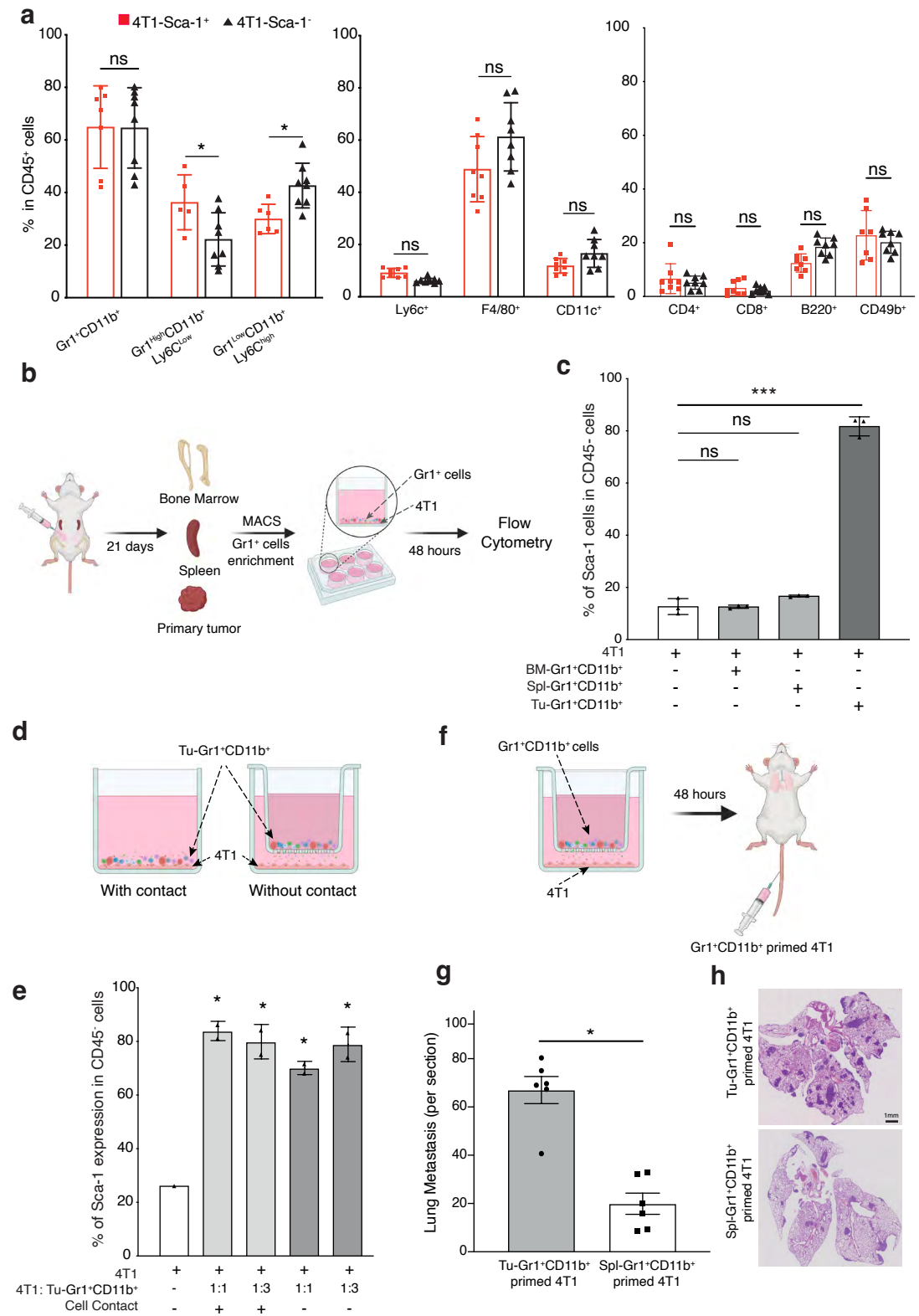

a
4T1-Sca-1+
4T1-Sca-1-
% in CD45+ cells
ns
*
Gr1+CD11b+
Gr1HighCD11b+ Ly6CLow
Gr1LowCD11b+ Ly6CHigh
Ly6c+
F4/80+
CD11c+
CD4+
CD8+
B220+
CD49b+
b
Bone Marrow
Spleen
Primary tumor
21 days
MACS
Gr1+ cells enrichment
Gr1+ cells
4T1
48 hours
Flow Cytometry
c
% of Sca-1 cells in CD45- cells
***
ns
4T1
BM-Gr1+CD11b+
Spl-Gr1+CD11b+
Tu-Gr1+CD11b+
d
Tu-Gr1+CD11b+
4T1
With contact
Without contact
f
Gr1+CD11b+ cells
48 hours
4T1
Gr1+CD11b+ primed 4T1
e
% of Sca-1 expression in CD45- cells
4T1
4T1: Tu-Gr1+CD11b+
Cell Contact
1:1
1:3
g
Lung Metastasis (per section)
Tu-Gr1+CD11b+ primed 4T1
Spl-Gr1+CD11b+ primed 4T1
h
Tu-Gr1+CD11b+ primed 4T1
Spl-Gr1+CD11b+ primed 4T1
1mm

Figure 2

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 2. Sca-1 expression is modulated by the tumor microenvironment

(a) Frequency of different immune cell populations in primary tumors of mice orthotopically injected with 4T1-Sca-1$^+$ and 4T1-Sca-1$^-$ 21 days post injection. Populations are determined in CD45 negative, viable cells.

(b-c) Illustrative scheme **(b)** showing the experimental design for isolating Gr1$^+$ cells from different sites of tumor-bearing mice. Twenty-one days after tumor implantation, Gr1$^+$ cells were isolated from bone marrow (BM-Gr1$^+$CD11b$^+$), spleen (Spl-Gr1$^+$CD11b$^+$) or primary tumor (Tu-Gr1$^+$CD11b$^+$) and co-cultured for 48 hours with parental 4T1 cells *in vitro.* Sca-1 expression in tumor cells was examined by flow cytometry **(c).** Co-cultures conditions are indicated in the bar graph.

(d) Illustrative scheme of the experimental co-culture setup.

(e) MACS-enriched Gr1$^+$ cells were cocultured with 4T1 cells with or without Transwell inserts of 0.4 µm pore size. The 4T1 cells were seeded in the bottom well and Gr1$^+$CD11b$^+$ cells in the upper part of the insert. After 48 hours, 4T1 cells were examined for Sca1 expression by FACS. Co-cultures conditions are indicated in the bar graph. The ratio of tumor cells and Tu-Gr1$^+$CD11b$^+$ is varied from 1:1 to 1:3. Significant enrichment of Sca-1$^+$ population were observed in all conditions.

(f) Illustrative scheme of the experimental metastasis setup.

(g-h) Evaluation of the metastatic capacity of Gr1$^+$CD11b$^+$-educated 4T1 cells *in vivo*. The 4T1 tumor cells were primed with Tu-Gr1$^+$CD11b$^+$ or Spl-Gr1$^+$CD11b$^+$ *in vitro* without cell-cell contact for 48 hours and injected into the tail vein of mice. Lung metastases were quantified 10 days after injection **(g)**, and representative images of lung sections are showed **(h).** n=7, H&E staining, Scale bar=1mm.

Data are represented as mean values ± SEM. P values: *, $p<0.05$; ***, $p< 0.001$ (unpaired two-tailed student's t test, and only significant data has been labeled).

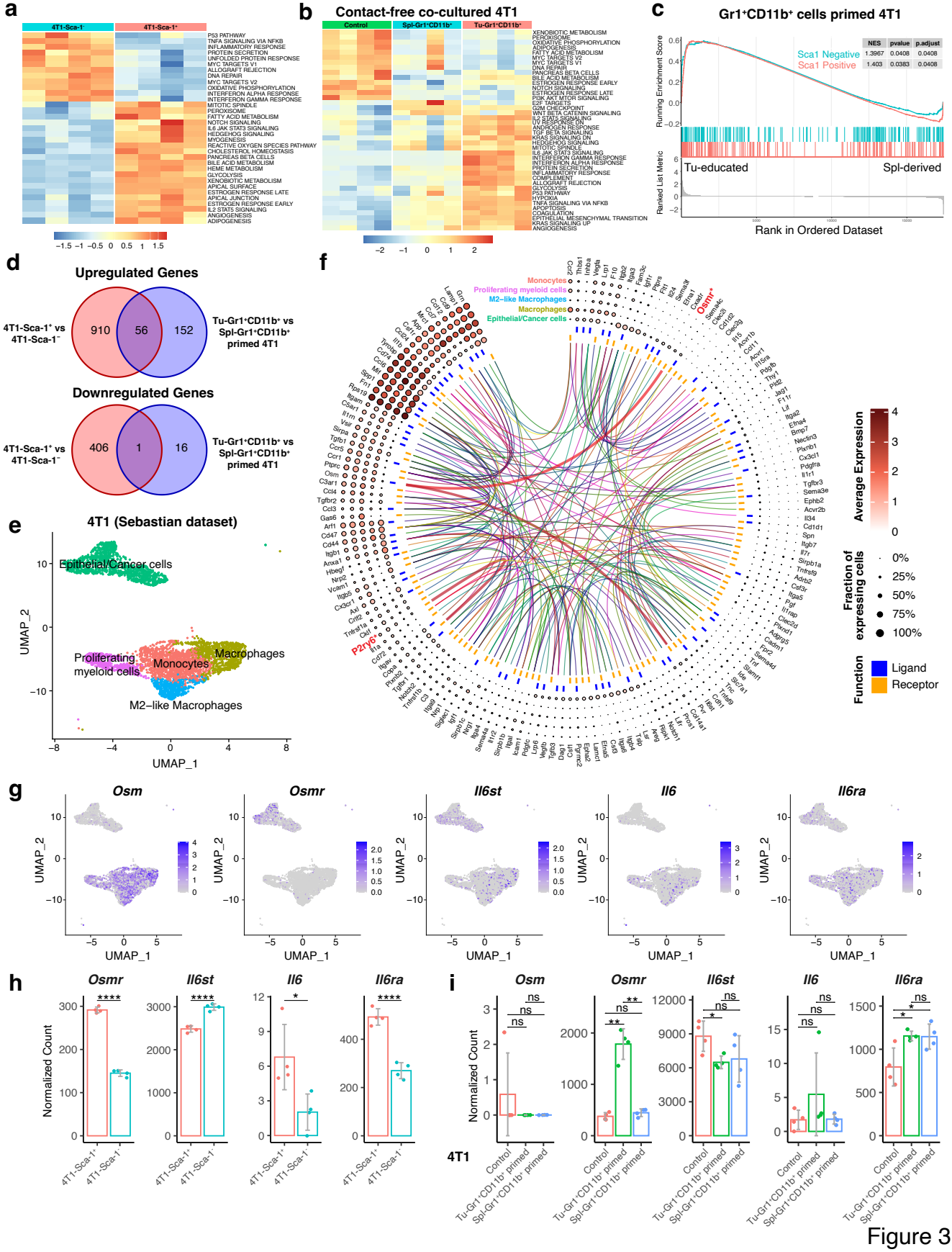

a
4T1-Sca-1⁻
4T1-Sca-1⁺
P53 PATHWAY
TNFA SIGNALING VIA NFKB
INFLAMMATORY RESPONSE
PROTEIN SECRETION
UNFOLDED PROTEIN RESPONSE
MYC TARGETS V1
ALLOGRAFT REJECTION
DNA REPAIR
MYC TARGETS V2
OXIDATIVE PHOSPHORYLATION
INTERFERON ALPHA RESPONSE
INTERFERON GAMMA RESPONSE
MITOTIC SPINDLE
PEROXISOME
FATTY ACID METABOLISM
NOTCH SIGNALING
IL6 JAK STAT3 SIGNALING
HEDGEHOG SIGNALING
MYOGENESIS
REACTIVE OXYGEN SPECIES PATHWAY
CHOLESTEROL HOMEOSTASIS
PANCREAS BETA CELLS
BILE ACID METABOLISM
HEME METABOLISM
GLYCOLYSIS
XENOBIOTIC METABOLISM
APICAL SURFACE
ESTROGEN RESPONSE LATE
APICAL JUNCTION
ESTROGEN RESPONSE EARLY
IL2 STAT5 SIGNALING
ANGIOGENESIS
ADIPOGENESIS
−1.5 −1 −0.5 0 0.5 1 1.5
b
Contact-free co-cultured 4T1
Control
Spl-Gr1⁺CD11b⁺
Tu-Gr1⁺CD11b⁺
XENOBIOTIC METABOLISM
PEROXISOME
OXIDATIVE PHOSPHORYLATION
ADIPOGENESIS
FATTY ACID METABOLISM
MYC TARGETS V2
MYC TARGETS V1
DNA REPAIR
PANCREAS BETA CELLS
BILE ACID METABOLISM
ESTROGEN RESPONSE EARLY
NOTCH SIGNALING
ESTROGEN RESPONSE LATE
PI3K AKT MTOR SIGNALING
E2F TARGETS
G2M CHECKPOINT
WNT BETA CATENIN SIGNALING
IL2 STAT5 SIGNALING
UV RESPONSE DN
ANDROGEN RESPONSE
TGF BETA SIGNALING
KRAS SIGNALING DN
HEDGEHOG SIGNALING
MITOTIC SPINDLE
IL6 JAK STAT3 SIGNALING
INTERFERON GAMMA RESPONSE
INTERFERON ALPHA RESPONSE
PROTEIN SECRETION
INFLAMMATORY RESPONSE
COMPLEMENT
ALLOGRAFT REJECTION
GLYCOLYSIS
P53 PATHWAY
HYPOXIA
TNFA SIGNALING VIA NFKB
APOPTOSIS
COAGULATION
EPITHELIAL MESENCHYMAL TRANSITION
KRAS SIGNALING UP
ANGIOGENESIS
−2 −1 0 1 2
c
Gr1⁺CD11b⁺ cells primed 4T1
NES pvalue p.adjust
Sca1 Negative 1.3967 0.0408 0.0408
Sca1 Positive 1.403 0.0383 0.0408
Running Enrichment Score
Ranked List Metric
Tu-educated
Spl-derived
Rank in Ordered Dataset
d
Upregulated Genes
4T1-Sca-1⁺ vs 4T1-Sca-1⁻
910
56
152
Tu-Gr1⁺CD11b⁺ vs Spl-Gr1⁺CD11b⁺ primed 4T1
Downregulated Genes
4T1-Sca-1⁺ vs 4T1-Sca-1⁻
406
1
16
Tu-Gr1⁺CD11b⁺ vs Spl-Gr1⁺CD11b⁺ primed 4T1
e
4T1 (Sebastian dataset)
Epithelial/Cancer cells
Proliferating myeloid cells
Monocytes
Macrophages
M2−like Macrophages
UMAP_1
UMAP_2
f
Monocytes
Proliferating myeloid cells
M2−like Macrophages
Macrophages
Epithelial/Cancer cells
Osmr⁺
P2rx7⁺
Average Expression
Fraction of expressing cells
0%
25%
50%
75%
100%
Function
Ligand
Receptor
g
Osm
Osmr
Il6st
Il6
Il6ra
UMAP_1
UMAP_2
h
Osmr
Il6st
Il6
Il6ra
Normalized Count
4T1-Sca-1⁺
4T1-Sca-1⁻
i
Osm
Osmr
Il6st
Il6
Il6ra
Normalized Count
4T1
Control
Tu-Gr1⁺CD11b⁺ primed
Spl-Gr1⁺CD11b⁺ primed

Figure 3

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 3. Transcriptomic analysis of Sca-1$^+$ tumor cells

(a) Heatmap showing the signature score of the hallmark pathways analysis in 4T1-Sca-1$^+$ and 4T1-Sca-1$^-$ population sorted from parental 4T1 cells. The colors code the expression levels relative to average levels as indicate at the bottom.

(b) Heatmap showing the signature score of the hallmarks pathway analysis in parental 4T1 (4T1), Spl-Gr1$^+$CD11b$^+$ primed 4T1 and Tu-Gr1$^+$CD11b$^+$ primed 4T1 cells. The colors code the expression levels relative to average levels as indicate at the bottom.

(c) Gene set enrichment analysis (GSEA) comparing the Tu-Gr1$^+$CD11b$^+$ and Spl-Gr1$^+$CD11b$^+$ primed 4T1 cells. GSEA shows positive correlations of both Sca1 Positive and Sca1 Negative signatures. NES, normalized enrichment score.

(d) Venn diagrams showing that 56 upregulated genes, and 1 downregulated gene are shared between endogenous and Tu-Gr1$^+$CD11b$^+$ induced Sca-1$^+$ population in 4T1 tumor cells.

(e) UMAP plot showing clusters of cancer cells and myeloid cell populations in orthotopically growing 4T1-derived primary tumors extracted from the Sebastian dataset (see Materials and Methods for details).

(f) Circos diagram showing the potential interactions between cancer cells and different myeloid cell populations determined by CellPhoneDB (see Materials and Methods for details) based on the Sebastian dataset. Only OSMR and P2RY6 are shared with the common 56 gene list showed in panel **d**.

(g) UMAP plot showing the gene expression pattern of *Osm*, *Osmr*, *Il6st*, *Il6* and *Il6ra* in different cell populations in the Sebastian dataset.

(h-i) mRNA expression of *Osm*, *Osmr*, *Il6st*, *Il6* and *Il6ra* based on RNAseq data used to generate the heatmaps shown in **a** & **b**, respectively. *Osm* expression is not detected in sorted 4T1 cells. Data are presented as mean of normalized count ± SD. P values:

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

*, $p<0.05$; **, $p< 0.01$; ***, $p< 0.001$; ****, $p< 0.0001$; ns, non-significant (unpaired two-tailed student's t test, and with Bonferroni correction for $Gr1^{+}CD11b^{+}$ cells-educated 4T1 cells).

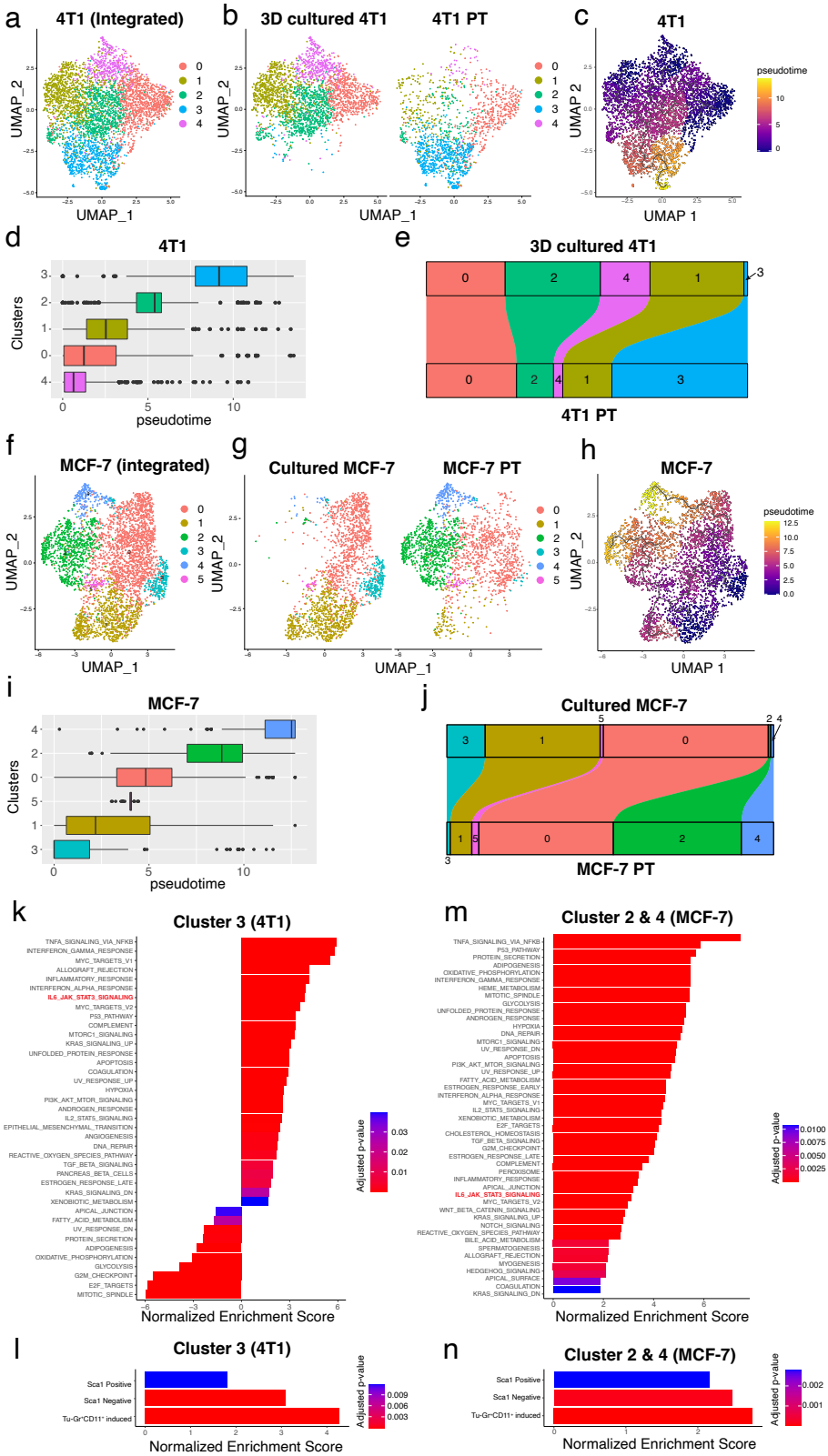

a
4T1 (Integrated)
UMAP_2
UMAP_1
0
1
2
3
4
b
3D cultured 4T1
4T1 PT
c
4T1
UMAP 1
pseudotime
d
4T1
Clusters
pseudotime
e
3D cultured 4T1
4T1 PT
f
MCF-7 (integrated)
g
Cultured MCF-7
MCF-7 PT
h
MCF-7
i
MCF-7
j
Cultured MCF-7
MCF-7 PT
k
Cluster 3 (4T1)
TNFA_SIGNALING_VIA_NFKB
INTERFERON_GAMMA_RESPONSE
MYC_TARGETS_V1
ALLOGRAFT_REJECTION
INFLAMMATORY_RESPONSE
INTERFERON_ALPHA_RESPONSE
IL6_JAK_STAT3_SIGNALING
MYC_TARGETS_V2
P53_PATHWAY
COMPLEMENT
MTORC1_SIGNALING
KRAS_SIGNALING_UP
UNFOLDED_PROTEIN_RESPONSE
APOPTOSIS
COAGULATION
UV_RESPONSE_UP
HYPOXIA
PI3K_AKT_MTOR_SIGNALING
ANDROGEN_RESPONSE
IL2_STAT5_SIGNALING
EPITHELIAL_MESENCHYMAL_TRANSITION
ANGIOGENESIS
DNA_REPAIR
REACTIVE_OXYGEN_SPECIES_PATHWAY
TGF_BETA_SIGNALING
PANCREAS_BETA_CELLS
ESTROGEN_RESPONSE_LATE
KRAS_SIGNALING_DN
XENOBIOTIC_METABOLISM
APICAL_JUNCTION
FATTY_ACID_METABOLISM
UV_RESPONSE_DN
PROTEIN_SECRETION
ADIPOGENESIS
OXIDATIVE_PHOSPHORYLATION
GLYCOLYSIS
G2M_CHECKPOINT
E2F_TARGETS
MITOTIC_SPINDLE
Adjusted p-value
Normalized Enrichment Score
m
Cluster 2 & 4 (MCF-7)
TNFA_SIGNALING_VIA_NFKB
P53_PATHWAY
PROTEIN_SECRETION
ADIPOGENESIS
OXIDATIVE_PHOSPHORYLATION
INTERFERON_GAMMA_RESPONSE
HEME_METABOLISM
MITOTIC_SPINDLE
GLYCOLYSIS
UNFOLDED_PROTEIN_RESPONSE
ANDROGEN_RESPONSE
HYPOXIA
DNA_REPAIR
MTORC1_SIGNALING
UV_RESPONSE_DN
APOPTOSIS
PI3K_AKT_MTOR_SIGNALING
UV_RESPONSE_UP
FATTY_ACID_METABOLISM
ESTROGEN_RESPONSE_EARLY
INTERFERON_ALPHA_RESPONSE
MYC_TARGETS_V1
IL2_STAT5_SIGNALING
XENOBIOTIC_METABOLISM
E2F_TARGETS
CHOLESTEROL_HOMEOSTASIS
TGF_BETA_SIGNALING
G2M_CHECKPOINT
ESTROGEN_RESPONSE_LATE
COMPLEMENT
PEROXISOME
INFLAMMATORY_RESPONSE
APICAL_JUNCTION
IL6_JAK_STAT3_SIGNALING
MYC_TARGETS_V2
WNT_BETA_CATENIN_SIGNALING
KRAS_SIGNALING_UP
NOTCH_SIGNALING
REACTIVE_OXYGEN_SPECIES_PATHWAY
BILE_ACID_METABOLISM
SPERMATOGENESIS
ALLOGRAFT_REJECTION
MYOGENESIS
HEDGEHOG_SIGNALING
APICAL_SURFACE
COAGULATION
KRAS_SIGNALING_DN
Adjusted p-value
Normalized Enrichment Score
l
Cluster 3 (4T1)
Sca1 Positive
Sca1 Negative
Tu-Gr-CD11+ induced
Adjusted p-value
Normalized Enrichment Score
n
Cluster 2 & 4 (MCF-7)
Sca1 Positive
Sca1 Negative
Tu-Gr-CD11+ induced
Adjusted p-value
Normalized Enrichment Score

Figure 4

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 4. Transformation dynamics of tumor cell populations induced by the tumor microenvironment

(a) UMAP plots showing 4T1 clusters based on integrated scRNA-seq data from 4T1 cells in 3D culture or in primary tumor.

(b) Distribution of specific clusters in 4T1 cells in 3D culture or in primary tumor (PT).

(**c-d)** UMAP plot **(c)** and boxplot **(d)** showing the clusters in pseudo-time during the transformation of 4T1 cells from *ex vivo* culture to *in vivo.*

(e) Sankey diagram showing the dynamic of each cluster during the transformation of 4T1 cells from *ex vivo* culture to *in vivo*. Cluster 3 was largely expanded *in vivo*.

(f) UMAP plots showing MCF-7 clusters based on integrated scRNA-seq data from MCF-7 cells in culture or in primary tumor.

(g) Distribution of the specific clusters in cultured MCF7 cells or in MCF7 primary tumors (PT).

(**h-i)** UMAP plot **(h)** and boxplot **(i)** showing the clusters in pseudo-time during the transformation of MCF-7 cells from *ex vivo* culture to *in vivo.*

(e) Sankey diagram showing the dynamic of each cluster during the transformation of MCF-7 cells from *ex vivo* culture to *in vivo*. Cluster 2 and 4 were largely expanded *in vivo*.

(k-l) GESA analysis of Hallmark gene sets **(k)** and Sca1 Positive signature, Sca1 Negative Signature and Tu-Gr1^{+}CD11b^{+} induced signature **(l)** of cluster 3 in 4T1 data. Only the signatures with adjusted p-value <0.05 were shown.

(m-n) GESA analysis of Hallmark gene sets **(m)** and Sca1 Positive signature, Sca1 Negative Signature and Tu-Gr1^{+}CD11b^{+} induced signature **(n)** of cells in cluster 2 or cluster 4 in MCF-7 data. Only the signatures with adjusted p-value <0.05 were shown.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Analyses are based on publicly available data (4T1: GSM4812003 and GSM3502134; MCF-7: GSM4681765 and GSM5904917).

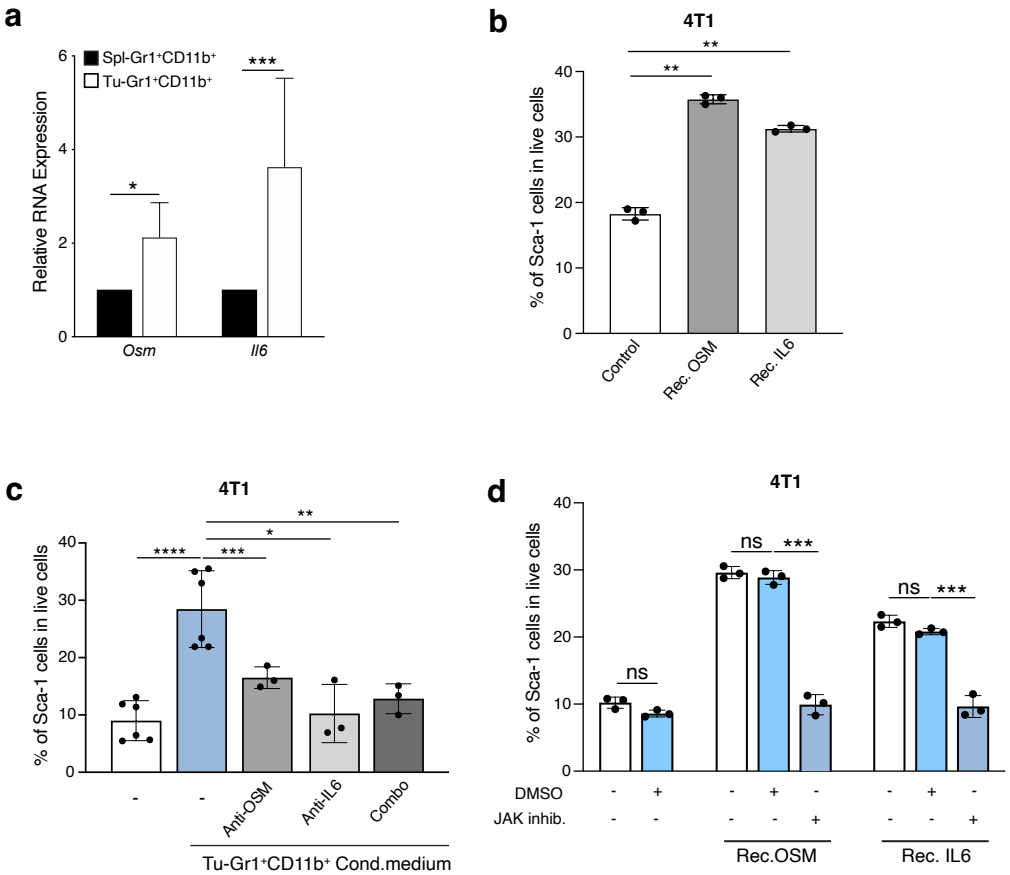

a
Spl-Gr1+CD11b+
Tu-Gr1+CD11b+
Relative RNA Expression
Osm
Il6
b
4T1
% of Sca-1 cells in live cells
Control
Rec. OSM
Rec. IL6
c
4T1
% of Sca-1 cells in live cells
Anti-OSM
Anti-IL6
Combo
Tu-Gr1+CD11b+ Cond.medium
d
4T1
% of Sca-1 cells in live cells
DMSO
JAK inhib.
Rec.OSM
Rec. IL6

Figure 5

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 5. Sca-1$^+$ population is modulated by the OSM/IL6-JAK pathway

(a) Quantitative PCR analysis of *Osm* and *Il6* mRNA expression in Tu-Gr1$^+$CD11b$^+$ and Spl- Gr1$^+$CD11b$^+$. Tu-Gr1$^+$CD11b$^+$ express significantly higher *Osm* and *Il6* levels compared with Spl-Gr1$^+$CD11b$^+$.

(b) Fraction of 4T1-Sca1$^+$ cells upon exposure to recombinant Il6 or Osm protein (1 µg/ml for 48 hours) as determined by flow cytometry. Both cytokines induced the Sca-1$^+$ population in cultured 4T1 tumor cells.

(c) Inhibition of OSM and IL6 from Tu-Gr1$^+$CD11b$^+$ conditioned medium with anti- OSM or anti-IL6 neutralizing antibody as indicated. Treatment with either antibody significantly suppressed the Sca-1$^+$ population enrichment.

(d) Treatment with the JAK inhibitor Ruxolitinib (5 µM) of cultured 4T1 cells stimulated with recombinant IL6 or OSM protein (1 µg/ml, 48 hours exposure) as indicated. Ruxolitinib inhibited Sca-1$^+$ population enrichment induced by recombinant IL6 or OSM protein.

Data are represented as mean ± SEM. *, $p<0.05$; **, $p< 0.01$; ***, $p< 0.001$; ****, $p< 0.0001$; ns, non-significant (unpaired two-tailed student's t test, and only significant data has been labeled).

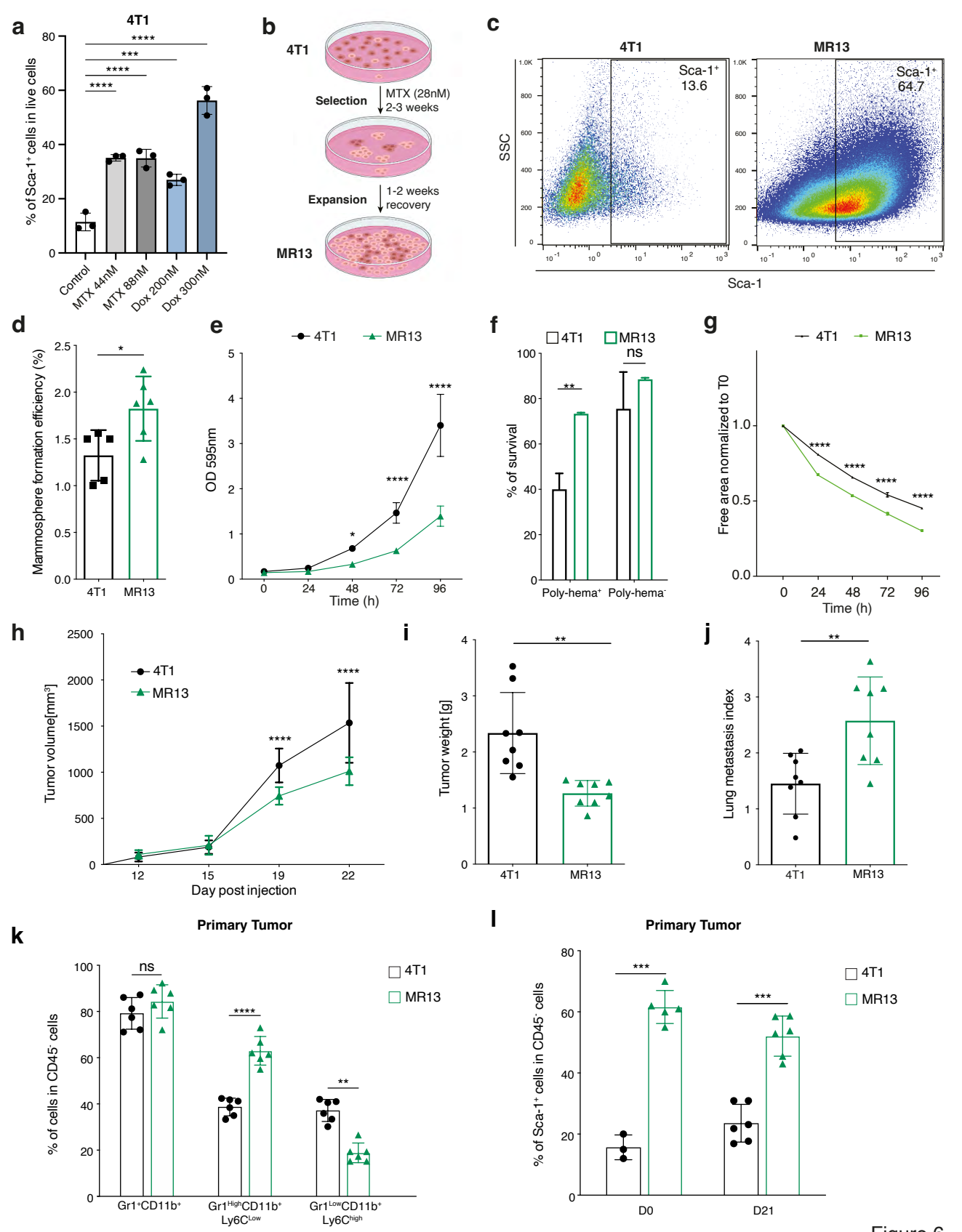

a
4T1
% of Sca-1+ cells in live cells
Control
MTX 44nM
MTX 88nM
Dox 200nM
Dox 300nM
b
4T1
Selection
MTX (28nM)
2-3 weeks
Expansion
1-2 weeks
recovery
MR13
c
4T1
MR13
Sca-1+
13.6
Sca-1+
64.7
SSC
Sca-1
d
Mammosphere formation efficiency (%)
4T1
MR13
e
4T1
MR13
OD 595nm
Time (h)
f
4T1
MR13
ns
% of survival
Poly-hema+
Poly-hema-
g
4T1
MR13
Free area normalized to T0
Time (h)
h
4T1
MR13
Tumor volume[mm3]
Day post injection
i
Tumor weight [g]
4T1
MR13
j
Lung metastasis index
4T1
MR13
k
Primary Tumor
ns
4T1
MR13
% of cells in CD45- cells
Gr1+CD11b+
Gr1HighCD11b+ Ly6CLow
Gr1LowCD11b+ Ly6Chigh
l
Primary Tumor
4T1
MR13
% of Sca-1+ cells in CD45- cells
D0
D21

Figure 6

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 6. Long term chemotherapy treatment of 4T1 cells induces a stable Sca-1$^+$ population (MR13) with higher metastatic capacity and CSC features

(a) Fraction of 4T1-Sca1$^+$ cells upon short-term (48 hours) Methotrexate (44 and 88 nM) and Doxorubicin (200 and 300 nM) treatments. All treatments induced enrichment of Sca-1$^+$ population.

(b) Illustrative scheme of the experimental design to obtain chemotherapy resistant MR13 cells from 4T1.

(c) Dot plots representing Sca-1 expression vs SSC determined by flow cytometry in MR13 chemotherapy resistant cells vs parental 4T1 tumor cells *in vitro.*

(d) Quantification of the mammosphere forming efficiency of 4T1 and MR13 tumor cells.

(e) Cell proliferation curve of 4T1 and MR13 tumor cells *in vitro* determined by crystal violet assay. The results are presented as mean of optical density (OD).

(f) Anchorage free survival of 4T1 and MR13 tumor cells given in % of surviving cells.

(g) Cell motility of 4T1 and MR13 tumor cells determined by a scratch wound healing assay. n=5-6/group. Results are presented as cell-free area relative to the initial wound area from 3 independent experiments.

(h) Growth curves of primary tumors in BALB/c mice orthotopically injected with 4T1 and MR13 tumor cells (n=10-11/group).

(i) Tumor weight of 4T1 and MR13 tumors recovered from BALB/c mice at day 22 post injection (n=8-9/group).

(j) Lung metastasis index 23 days post injection. The number of metastatic nodules is determined by H&E staining and normalized based on the primary tumor weight (n=8-9/group).

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(k) Frequency of different CD11b$^+$ myeloid cells subpopulations in primary tumors from MR13 and 4T1 injected mice determined by flow cytometry 21-days post injection (n= 6). Subpopulations are determined in CD45 positive, viable cell population.

(l) Percentage of Sca-1$^+$ tumor cells at time of injections (D0) of 4T1 and MR13 cells and in primary tumors recovered at day 21 (D21). Sca-1 expression is determined in CD45 negative, viable cell population.

Data are represented as mean values ± SEM. *, $p<0.05$; **, $p<0.005$; ***, $p< 0.001$, ****, $p< 0.0001$ unpaired two-tailed student's t test, and only significant data has been labeled. In e, g, h, results were analyzed by two-way ANOVA with Tukey's multiple comparison test.

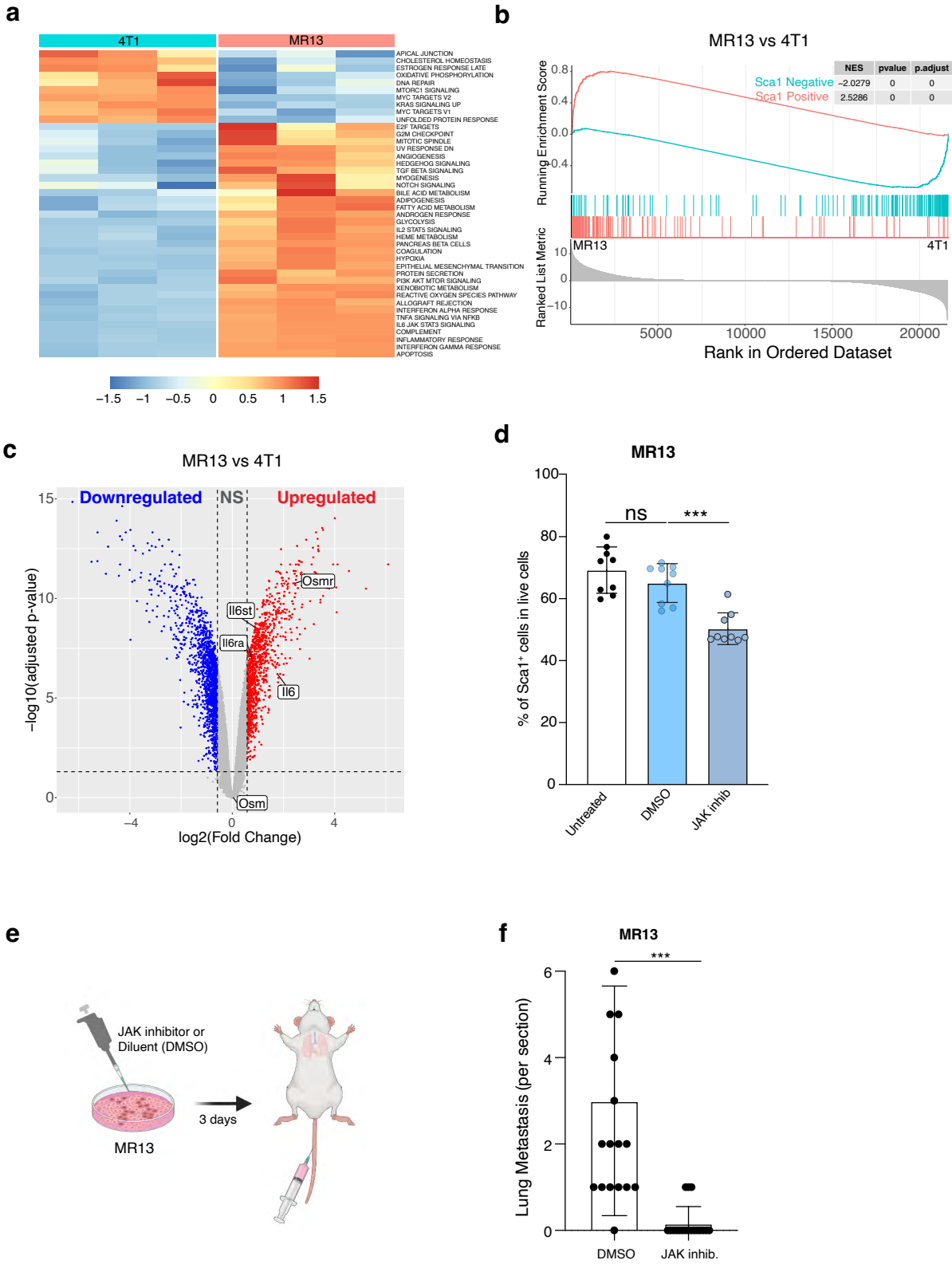

a
4T1
MR13
APICAL JUNCTION
CHOLESTEROL HOMEOSTASIS
ESTROGEN RESPONSE LATE
OXIDATIVE PHOSPHORYLATION
DNA REPAIR
MTORC1 SIGNALING
MYC TARGETS V2
KRAS SIGNALING UP
MYC TARGETS V1
UNFOLDED PROTEIN RESPONSE
E2F TARGETS
G2M CHECKPOINT
MITOTIC SPINDLE
UV RESPONSE DN
ANGIOGENESIS
HEDGEHOG SIGNALING
TGF BETA SIGNALING
MYOGENESIS
NOTCH SIGNALING
BILE ACID METABOLISM
ADIPOGENESIS
FATTY ACID METABOLISM
ANDROGEN RESPONSE
GLYCOLYSIS
IL2 STAT5 SIGNALING
HEME METABOLISM
PANCREAS BETA CELLS
COAGULATION
HYPOXIA
EPITHELIAL MESENCHYMAL TRANSITION
PROTEIN SECRETION
PI3K AKT MTOR SIGNALING
XENOBIOTIC METABOLISM
REACTIVE OXYGEN SPECIES PATHWAY
ALLOGRAFT REJECTION
INTERFERON ALPHA RESPONSE
TNFA SIGNALING VIA NFKB
IL6 JAK STAT3 SIGNALING
COMPLEMENT
INFLAMMATORY RESPONSE
INTERFERON GAMMA RESPONSE
APOPTOSIS
−1.5 −1 −0.5 0 0.5 1 1.5
b
MR13 vs 4T1
NES pvalue p.adjust
Sca1 Negative −2.0279 0 0
Sca1 Positive 2.5286 0 0
Running Enrichment Score
Ranked List Metric
MR13
4T1
Rank in Ordered Dataset
c
MR13 vs 4T1
Downregulated
NS
Upregulated
Osmr
Il6st
Il6ra
Il6
Osm
−log10(adjusted p-value)
log2(Fold Change)
d
MR13
ns
***
% of Sca1+ cells in live cells
Untreated
DMSO
JAK inhib
e
JAK inhibitor or
Diluent (DMSO)
MR13
3 days
f
MR13
***
Lung Metastasis (per section)
DMSO
JAK inhib.

Figure 7

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 7. IL6-JAK pathway promotes Sca-1$^+$ persistence and metastatic capacity in chemotherapy resistant MR13 tumor cells

(a-b) Gene expression analysis of parental 4T1 and chemotherapy resistant MR13 cells. Heat map represents the signature score of the hallmark pathways analysis. Results from 3 biological replicates are shown **(a)**. GSEA results showing that MR13 cells are positively enriched with for the Sca-1 Positive signature and negatively with for the Sca-1 Negative signature **(b)**.

(c) Volcano plot showing the differential expression of *Osm*, *Osmr*, *Il6st*, *Il6* and *Il6ra* mRNA in MR13 vs 4T1 tumor cells.

(d) Fraction of Sca-1$^+$ population in MR13 tumor cells treated for 48 hours with Ruxolitinib (5 µM) relative to vehicle control (DMSO) treatment.

(e) Illustrative scheme of the experimental design for testing the effect of Ruxolitinib on MR13 metastatic capacity shown in **f**. MR13 tumor cells were treated with Ruxolitinib or DMSO *in vitro* for 72 hours and then injected into the mice tail vein. Lungs were examined for metastasis 10 days after tumor cell injection.

(f) Number of metastatic nodules in the lungs from mice injected with MR13 treated *in vitro* with Ruxolitinib or DMSO and indicated (n=8).

Data are represented as mean values ± SEM. P values: ns, non-significant, ***< 0.001, (unpaired two-tailed student's t test, and with Bonferroni correction for d)

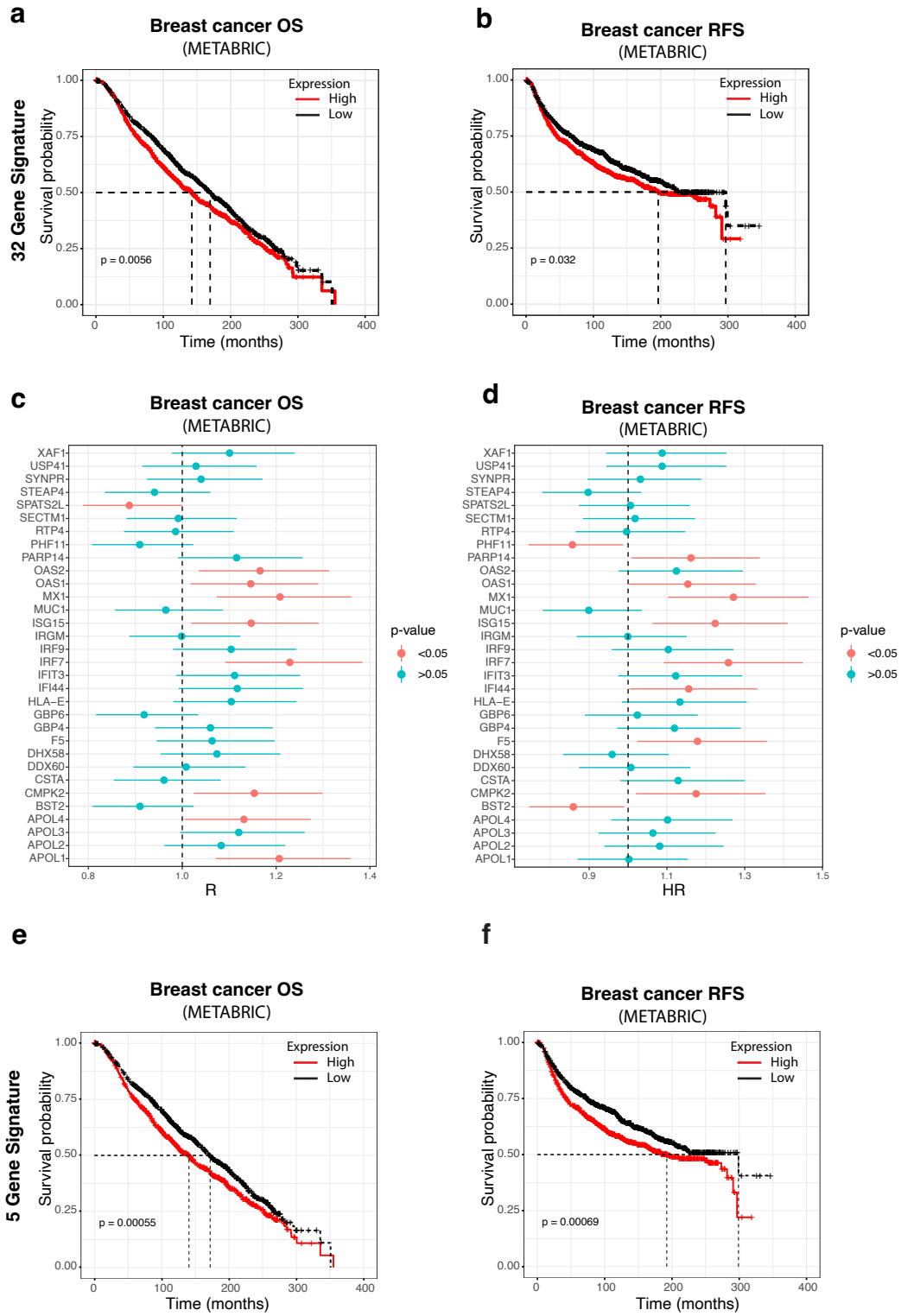

a
Breast cancer OS
(METABRIC)
32 Gene Signature
Survival probability
Expression
High
Low
p = 0.0056
Time (months)
b
Breast cancer RFS
(METABRIC)
Survival probability
Expression
High
Low
p = 0.032
Time (months)
c
Breast cancer OS
(METABRIC)
XAF1
USP41
SYNPR
STEAP4
SPATS2L
SECTM1
RTP4
PHF11
PARP14
OAS2
OAS1
MX1
MUC1
ISG15
IRGM
IRF9
IRF7
IFIT3
IFI44
HLA−E
GBP6
GBP4
F5
DHX58
DDX60
CSTA
CMPK2
BST2
APOL4
APOL3
APOL2
APOL1
R
p-value
<0.05
>0.05
d
Breast cancer RFS
(METABRIC)
HR
p-value
<0.05
>0.05
e
Breast cancer OS
(METABRIC)
5 Gene Signature
Survival probability
Expression
High
Low
p = 0.00055
Time (months)
f
Breast cancer RFS
(METABRIC)
Survival probability
Expression
High
Low
p = 0.00069
Time (months)

Figure 8

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 8. Tu-Gr1$^+$CD11b$^+$-induced tumor cell signature predicts worse outcome in breast cancer patients

(a-b) Kaplan-Meier curves showing overall survival (OS) **(a)** or relapse-free survival (RFS) **(b)** for breast cancer patients according to high or low expression of an orthologue 32 gene signature, based on the Tu-Gr1$^+$CD11b$^+$ -induced 4T1 cell signature, in the METABRIC datasets. The p-value was calculated using the log-rank test and high and low expression levels were stratified by median values.

(c-d) Forest plots showing the Cox proportional hazard regression (HR) for OS **(c)** and RFS **(d)** of the individual 32 orthologues of the Tu-Gr1$^+$CD11b$^+$-induced signature, based on gene expression in tumor samples from METABRIC dataset.

(e-f) Kaplan-Meier curves showing OS **(e)** and RFS **(f)** according to the reduced 5 orthologue gene signature. expression in the METABRIC datasets.

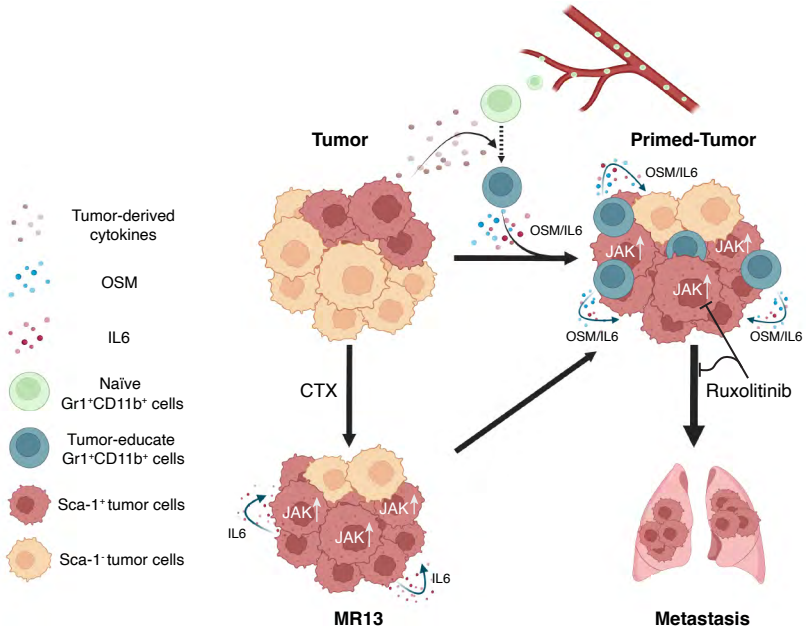
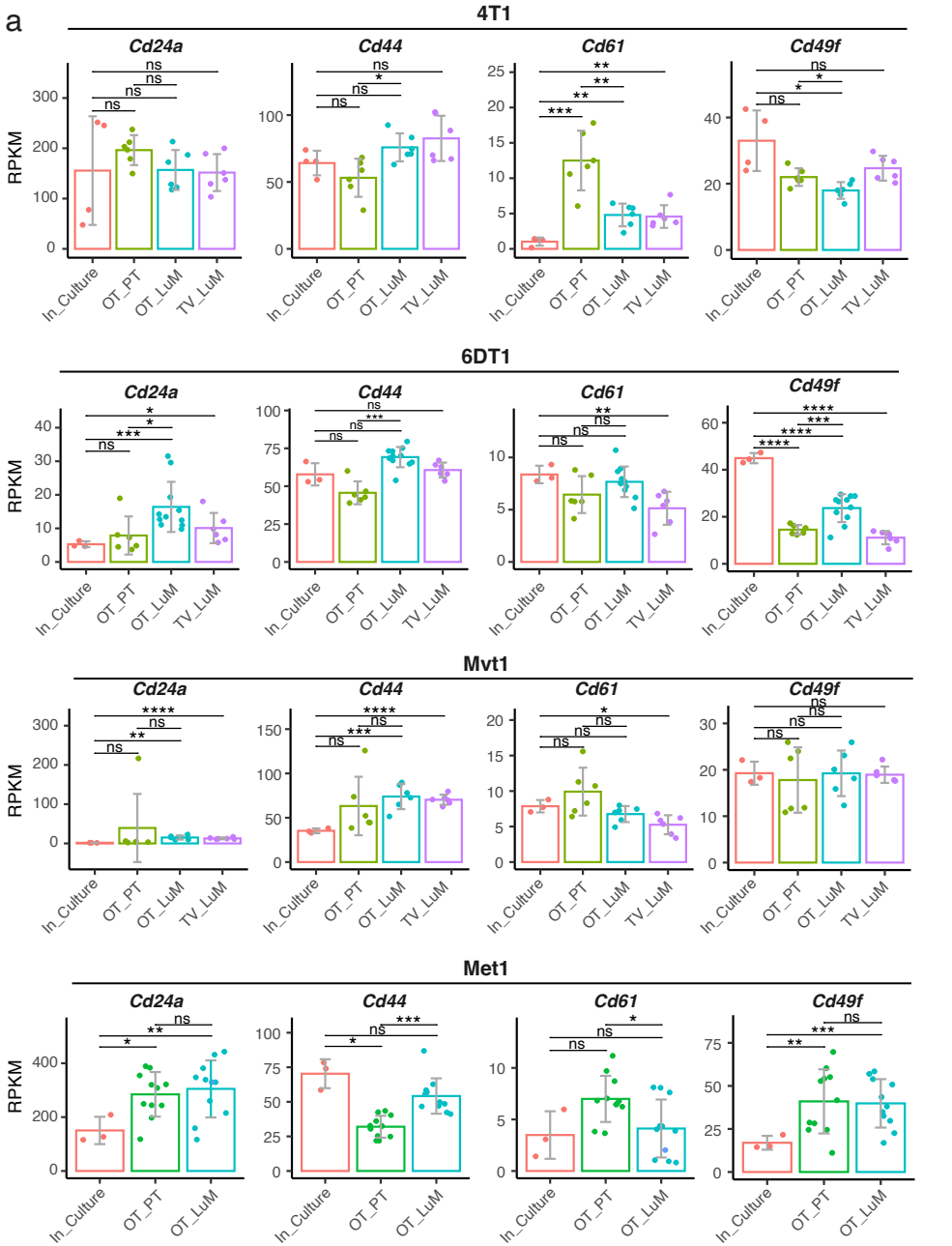

Tumor-derived cytokines
OSM
IL6
Naïve Gr1+CD11b+ cells
Tumor-educate Gr1+CD11b+ cells
Sca-1+ tumor cells
Sca-1- tumor cells
Tumor
Primed-Tumor
OSM/IL6
OSM/IL6
JAK
JAK
JAK
OSM/IL6
OSM/IL6
CTX
Ruxolitinib
JAK
JAK
JAK
IL6
IL6
MR13
Metastasis

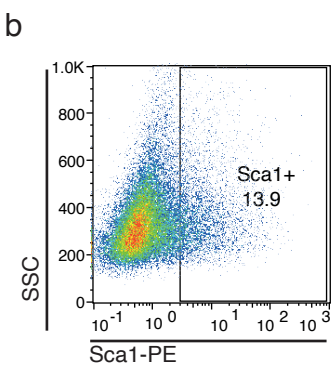
Figure 9

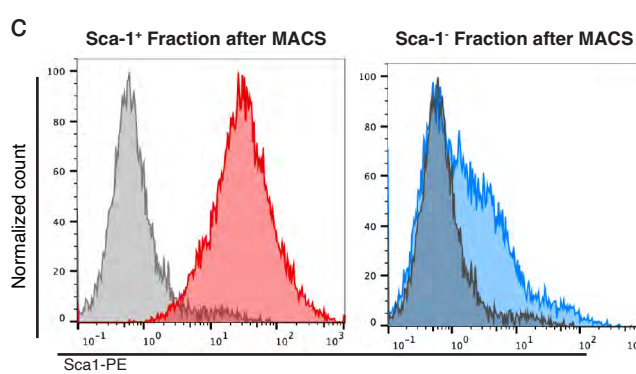
bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 9 Illustrative scheme of the proposed model for cancer cell plasticity modulated by OSM/IL6 during tumor progression and chemotherapy

Parental tumor cells maintain a small portion of highly metastatic Sca-1$^+$ population. During tumor progression, naïve Gr1$^+$CD11b$^+$ are recruited to the TME and educated into Tu-Gr1$^+$CD11b$^+$ by tumor-derived factors. In turn, Tu-Gr1$^+$CD11b$^+$ secrete OSM and IL6 to convert Sca-1$^-$ population into a highly metastatic Sca-1$^+$ population. Chemotherapy (CTX) enriched for Sca-1$^+$ population due to its intrinsic resistance against cytotoxic treatment. Resistant cells express IL6 to maintain the high portion of Sca-1$^+$ population with high metastatic ability. JAK inhibitor Ruxolitinib suppresses the conversion to Sca-1$^+$ population and metastasis. Inhibition of OSM, IL6 and the activated downstream kinase JAK are candidate therapeutic targets to impinge on metastatic breast cancer progression during natural evolution and following therapy-resistance.







Supplementary Fig.1

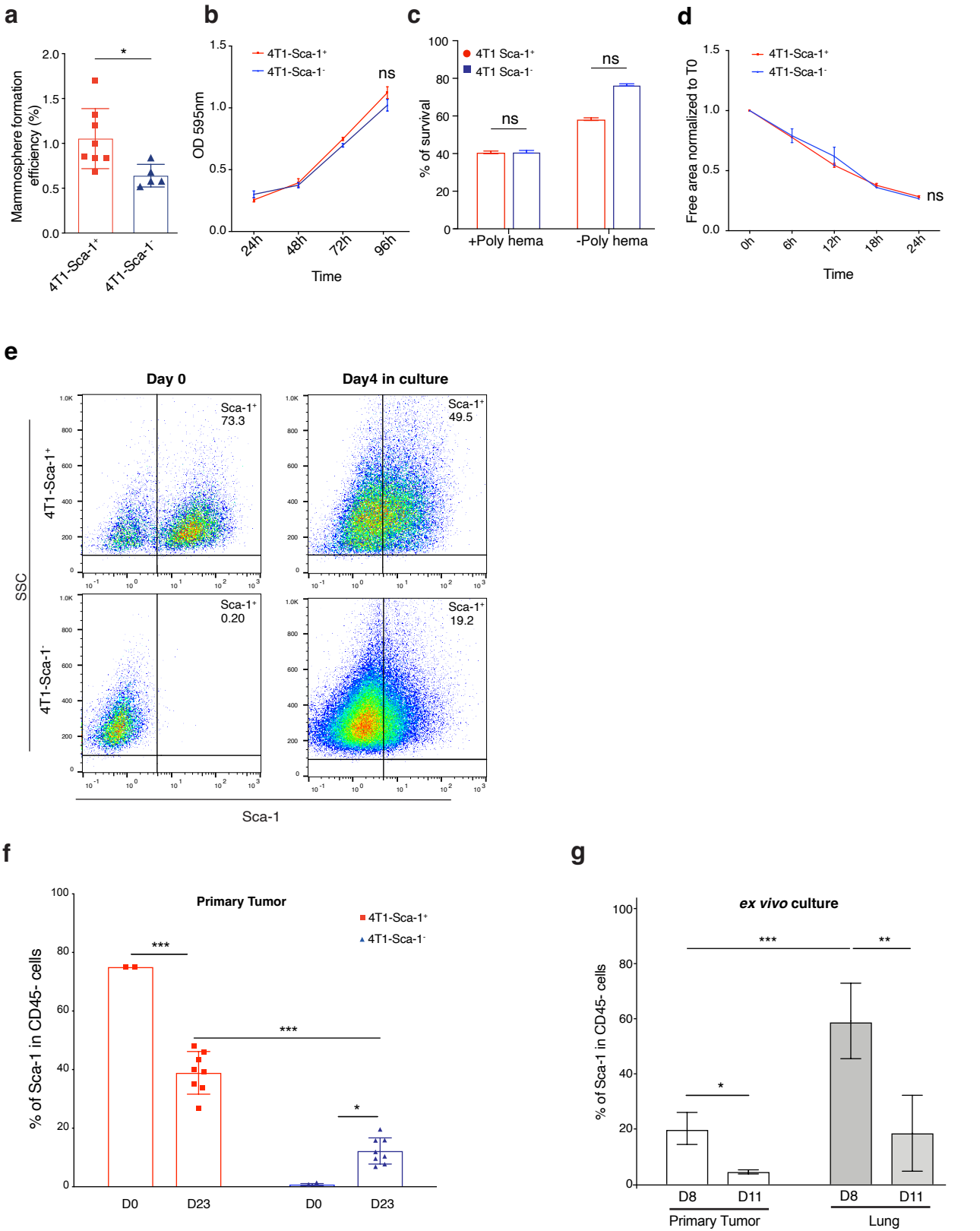
bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplementary Fig 1. Stem cell markers expression in multiple metastatic breast cancer models

(a) Stem cell marker *Cd24a*, *Cd44*, *Cd61* and *Cd49e* mRNA expression in 4T1, 6DT1, Mvt1 and Met1 metastatic murine breast cancer models extracted from Ross dataset. Data are presented as mean values of RPKM ± SD. *, $p<0.05$; **, $p< 0.01$; ***, $p< 0.001$; ****, $p< 0.0001$; ns, non-significant (unpaired two-tailed student's t test).

(b) Dot plot representation of Sca-1 expression in 4T1 parental cell line determined by flow cytometry.

(c) Histogram of Sca-1 expression distribution on MACS positively selected Sca-1^{+} cells (red histogram) and MACS negatively selected Sca-1^{-} cell (blue histogram). Gray: histogram of fluorescence of unstained cells.



Supplementary Fig. 2

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplementary Fig 2. Sca-1$^+$ tumor cells have stem cell-like features *in vitro* and show plasticity *in vivo* and *in vitro*

(a) Quantification of the mammosphere forming efficiency of 4T1-Sca-1$^+$ and 4T1-Sca-1$^-$ populations.

(b) Cell proliferation curve of 4T1-Sca-1$^+$ and 4T1-Sca-1$^-$ *in vitro* MACS isolated tumor cells determined by crystal violet assay. The results represent optical density (OD) of the wells.

(c) Anchorage-independent survival of 4T1-Sca-1$^+$ and 4T1-Sca-1$^-$ MACS isolated tumor cells determined with an anoikis assay. Cells were cultured in control or poly-hema coated wells (to avoid adhesion) and viability measured by annexin/PI staining and FACS analysis. The results are shown as % of viable cells identified as Annexin V negative and PI negative cells.

(d) Cell motility of 4T1-Sca-1$^+$ and 4T1-Sca-1$^-$ sorted tumor cells determined by a scratch wound healing assay (n=5-6/group). Results are given cell-free area relative to the initial wound area.

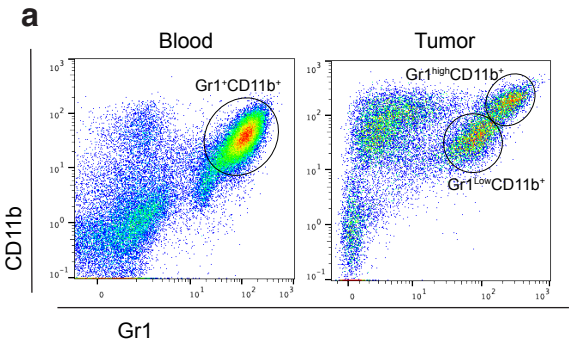
(e) The abundance of Sca-1$^+$ population after MACS isolation from parental 4T1 cells. Day 0: immediately after positive sorting, the Sca-1$^+$ population to accounts for 73.3% of total cells (upper panel). After 4 days of *in vitro* culture the abundance of the Sca-1$^+$ population decreased to 48.9%. Negatively sorted Sca-1$^-$ population accounts for >99 % of total cells (lower panel). After 4 days of *in vitro* culture 19.2% of the initially Sca-1$^+$ cells were Sca-1$^+$.

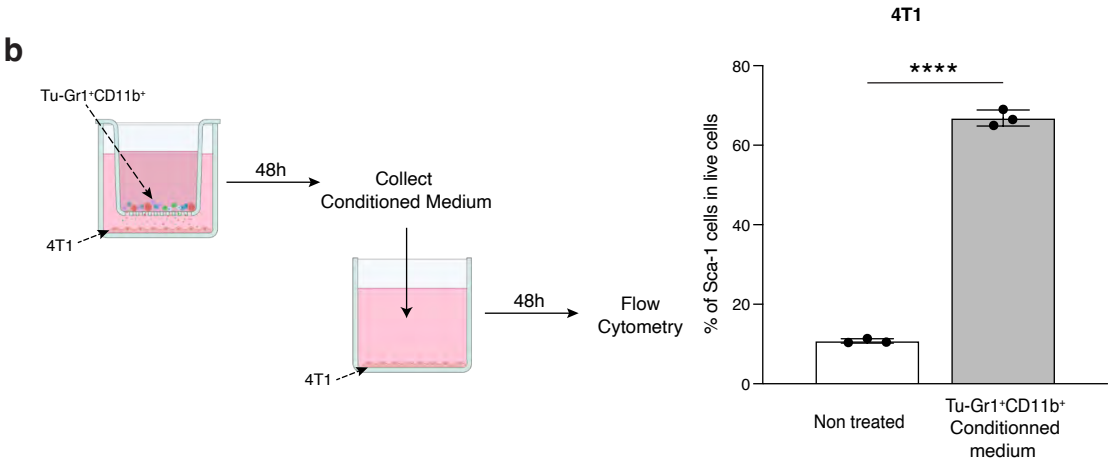
(f) Abundance of Sca-1$^+$ population at the time of orthotopic injection of MACS isolated 4T1-Sca-1$^+$ and 4T1-Sca-1$^-$ cells into the mammary fat pad of BALB/C mice (D0) and in the derived primary tumors 23 days post injection (D23) as indicated. n=3-8/group.

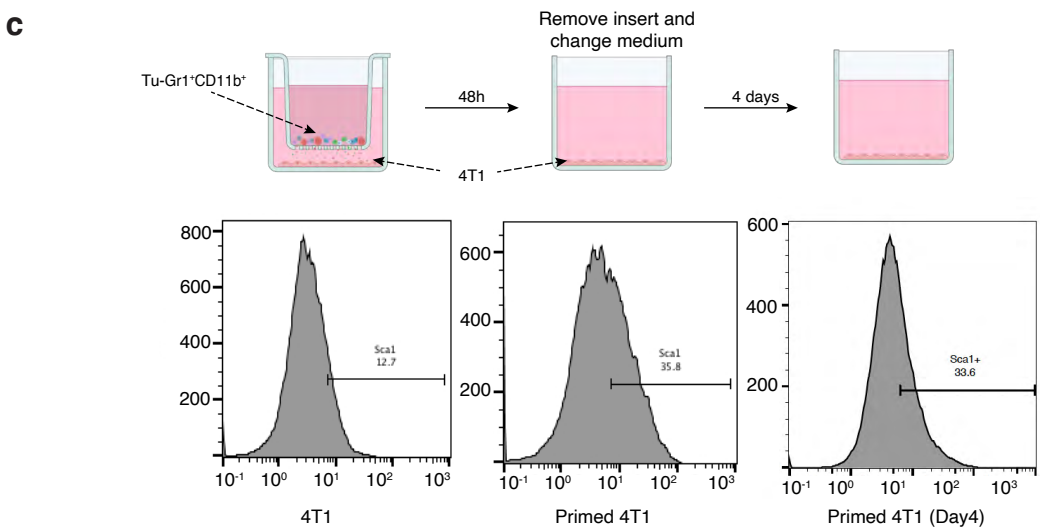
bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(g) Abundancy of the Sca-1^{+} population in tumor cells recovered from primary tumors and lung metastases 21 days post orthotropic injection and further cultured for 8- and 11-days *ex vivo* as indicated.

Data represent mean values ± SEM from 3 independent experiments. *, $p<0.05$; **, $p< 0.01$; ***, $p< 0.001$; (unpaired two-tailed student's t test).







Supplementary Fig. 3

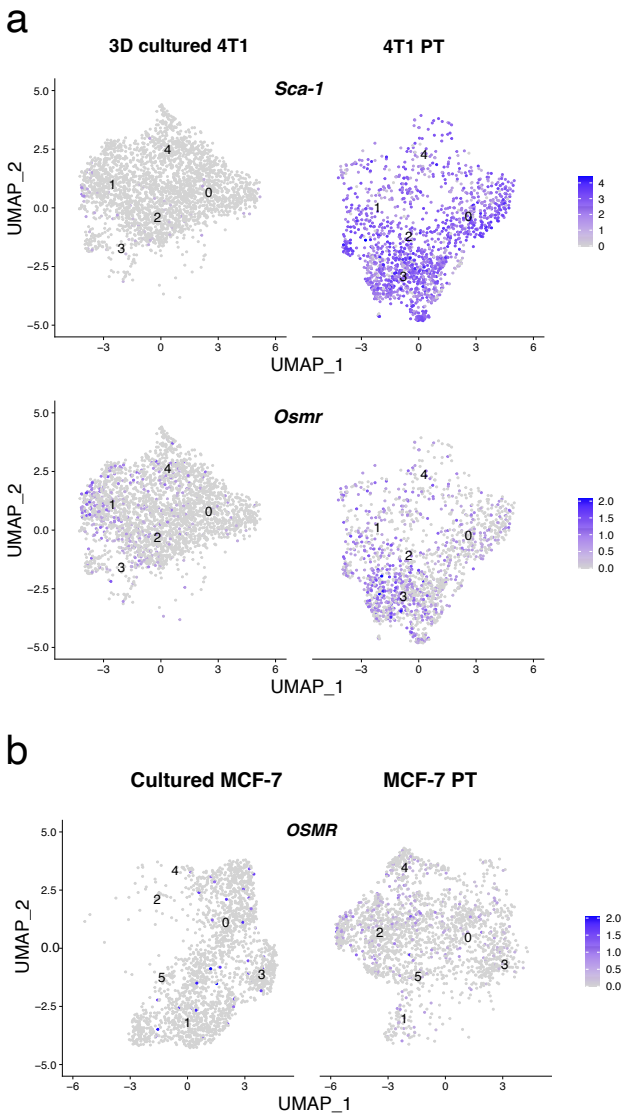
bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplementary Fig 3. Tu-Gr1$^+$CD11b$^+$-derived secreted factors promote the enrichment of Sca-1$^+$ population

(a) Dot plots presenting Gr1$^+$CD11b$^+$ subpopulations in Blood and Tumor site of 4T1 tumor-bearing mice.

(b) Abundance of the Sca-1$^+$ population in tumor cells cultured for 48 hours in the presence of medium collected from 4T1/Tu-Gr1$^+$CD11b$^+$ cocultures (48 hours). Sca-1 expression on 4T1 cells was determined by flow cytometry.

(c) Sca-1 expression in parental 4T1 cells (left panel), 4T1 primed for 2 days with Gr1$^+$CD11b$^+$ cells (middle panel) and cultured for 4 additional days in the absence of Gr1$^+$CD11b$^+$ (right panel). Sca-1 expression was determined by flow cytometry.



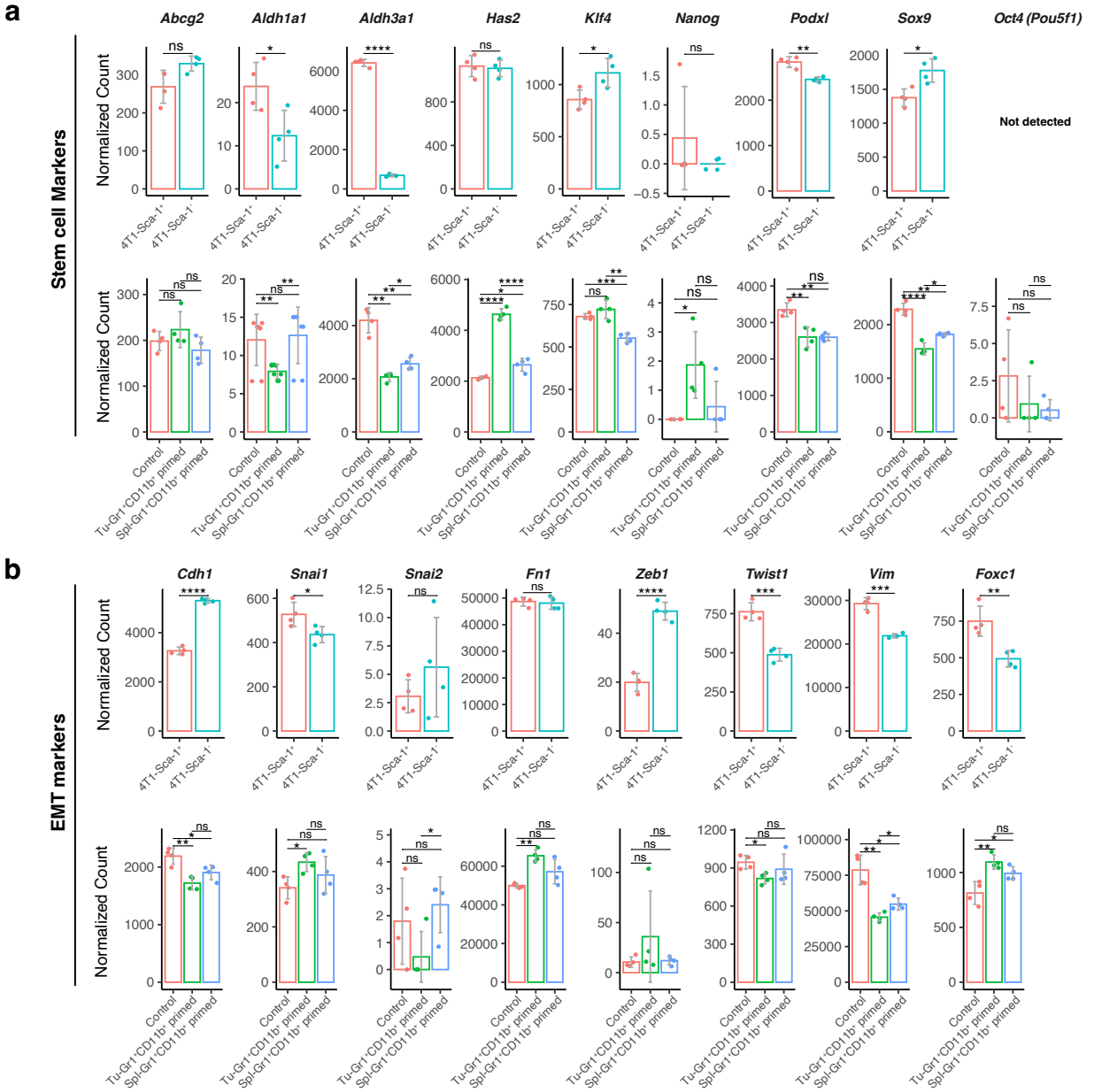
Supplementary Fig. 4

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplementary Fig 4. TME expands the Sca1^{+} population during *in vivo* tumor progression

(a) UMAP plots showing the expression of *Sca-1* and *Osmr* in 4T1 tumor cells in 3D culture or primary tumors (PT). Analysis based on publicly available data (GSM4812003 and GSM3502134)

(b) UMAP plot showing the expression of *OSMR* in MCF-7 tumor cells in culture or primary tumors (PT). Analysis based on publicly available data (GSM4681765 and GSM5904917).



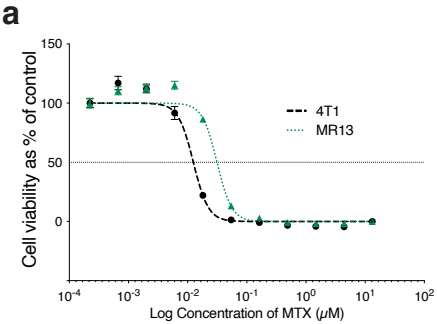
Supplementary Fig. 5

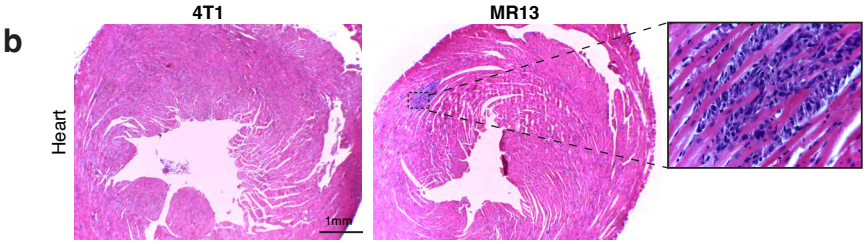
bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplementary Fig 5. Expression of selected stem cell/CSC and EMT genes in 4T1-**Sca-1$^+$,** 4T1-**Sca-1$^-$ and differently primed 4T1 cells**

Expression of stem cell markers (**a**) and EMT markers (**b**) in 4T1-Sca-1$^+$ cells, 4T1-Sca-1$^-$ cells, and parental, Tu-Gr1$^+$CD11b$^+$ -educated, Spl-Gr1$^+$CD11b$^+$ -educated 4T1 cells, as indicated.

Data represent mean values of RPKM ± SD. *, $p<0.05$; **, $p< 0.01$; ***, $p< 0.001$; ****, $p< 0.0001$ (unpaired two-tailed student's t test, and with Bonferroni correction for MDSC-educated 4T1 cells).





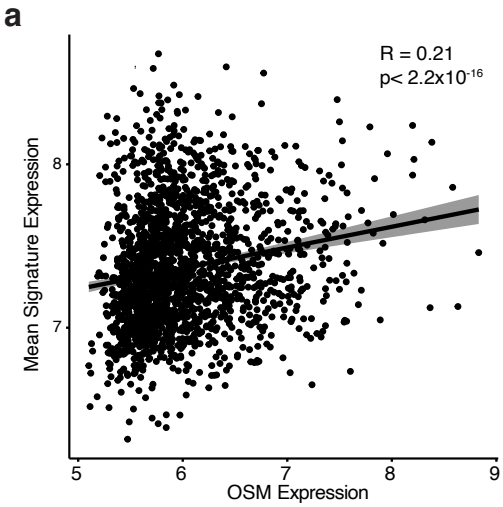
Supplementary Fig. 6

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplementary Fig 6. Sca-1$^+$ enriched 4T1 cells and MR13 cells are more resistant to chemotherapy drugs

(a) Dose dependent effect of 48 hours treatment with MTX on 4T1 and MR13 cells' viability. IC_{50} of MTX for 4T1 and MR13 are 12.5 nM and 31.4 nM respectively.

(b) Illustrative H&E stained heart sections of BALB/c mice 22 days post orthotopic injection with 4T1 cells (left) or MR13 cells (right and insert) (n=8-9/group).



Supplementary Fig. 7

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.26.513889; this version posted October 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplementary Fig 7. Genes of the Tu-Gr1$^+$CD11b$^+$-induced signature and patients' outcome

(a) Correlation of the orthologue 32 gene signature, based on the Tu-Gr1$^+$CD11b$^+$ -induced signature, with *OSM* expression in the METABRIC datasets. Spearman's correlation coefficients and p-values are shown.